

Přírodovědecká fakulta
Katedra geoinformatiky

**VYMEZENÍ VÝTVARNÉHO STYLU
TEMATICKÝCH MAP
METODAMI MACHINE LEARNING**

diplomová práce

Bc. Martin SADÍLEK

Vedoucí práce: prof. RNDr. Vít VOŽENÍLEK, CSc.

Olomouc 2021
Geoinformatika

ANOTACE

Výtvarný styl je jednou z prvních věcí, které jsou v mapách nepřímo hodnoceny téměř každým uživatelem. Současná literatura však nenabízí žádnou společně dohodnutou definici výtvarného stylu, ani nerozlišuje mezi jednotlivými výtvarnými styly. Tato diplomová práce se zabývá možnostmi metod strojového učení, konkrétně neuronových sítí, k identifikaci těchto stylů u tematických map. Předpoklad pro práci je ten, že neuronové sítě mohou výrazně přispět k definici výtvarných stylů map. Doposud sice neexistuje žádná podobná studie, která by zkoumala definici výtvarného stylu obrazových dat, ke kterým mapy patří, natož s použitím softwaru Orange. Důvodem bude pravděpodobně jen nevyužití možností strojového učení v této oblasti, protože v jiných odvětvích, jako jsou IT, přírodní vědy nebo technické vědy, jsou tyto nástroje široce používány. Neuronové sítě z tohoto softwaru již byly nicméně úspěšně použity k analýze podobnosti kruhových výřezů měst z Urban Atlasu (Dobešová 2019). Tato diplomová práce je proto v mnoha ohledech průlomová a poskytuje náhled na potenciál metod machine learning (konkrétně neuronových sítí) a využití softwaru Orange k ještě nerealizovaným kartografickým úlohám.

KLÍČOVÁ SLOVA

výtvarný styl; machine learning; ORANGE

Počet stran práce: 68

Počet příloh: 10 (z toho 7 volné, 2 vázané a 1 elektronická)

ANOTATION

Artistic style is one of the first things that are indirectly evaluated in maps by almost every user. However, current literature offers no commonly agreed upon definition of an artistic style nor does it differentiate between artistic styles. This project work deals with the application of the machine learning methods, specifically neural networks, for the identification of an artistic style of a map. The assumption is that neural networks can contribute to a definition of a map artistic style because neural networks were already successfully used to analyse similarities in pattern of circular sections of cities from the Urban Atlas as reported by Dobešová (2019). Until now, there is no similar study examining the artistic style definition of the image data, to which the maps belong, let alone using Orange software for study like this, an open-source software created at the University of Ljubljana in Slovenia (Orange Data Mining, 2020). The reason for that will probably be just the non-use of machine learning opportunities in this area, because in other sectors like IT, natural sciences or technical sciences, these tools are used widely. This work is therefore ground-breaking in many respects and provides insight into the use of the possibilities of this software for yet unrealized tasks in this area.

KEYWORDS

artistic style; machine learning; ORANGE

Number of pages: 68

Number of appendixes: 10

Prohlašuji, že

- diplomovou práci včetně příloh, jsem vypracoval samostatně a uvedl jsem všechny použité podklady a literaturu,

- jsem si vědom, že na moji diplomovou práci se plně vztahuje zákon č.121/2000 Sb. - autorský zákon, zejména § 35 – využití díla v rámci občanských a náboženských obřadů, v rámci školních představení a využití díla školního a § 60 – školní dílo,

- beru na vědomí, že Univerzita Palackého v Olomouci (dále UP Olomouc) má právo nevydělečně, ke své vnitřní potřebě, diplomovou práci užívat (§ 35 odst. 3),

- souhlasím, aby jeden výtisk diplomové práce byl uložen v Knihovně UP k prezenčnímu nahlédnutí,

- souhlasím, že údaje o mé diplomové práci budou zveřejněny ve Studijním informačním systému UP,

- v případě zájmu UP Olomouc uzavřu licenční smlouvu s oprávněním užít výsledky a výstupy mé diplomové práce v rozsahu § 12 odst. 4 autorského zákona,

- použít výsledky a výstupy mé diplomové práce nebo poskytnout licenci k jejímu využití mohu jen se souhlasem UP Olomouc, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly UP Olomouc na vytvoření díla vynaloženy (až do jejich skutečné výše).

V Olomouci dne

Bc. Martin Sadílek

Rád bych poděkoval vedoucímu své diplomové práce prof. RNDr. Vítu Voženilkovi, CSc. za odborné vedení, za pomoc a rady při zpracování této práce. Poděkování patří také mé rodině, za psychickou a materiální podporu. Za to, že mi umožnili studium a vždy mě plně podporovali. Poděkování patří také mým kamarádům a spolužákům Elišce Regentové, Petrovi Mužičkovi a Tereze Novákové, za oporu a zpříjemnění studijních let. Velké poděkování patří mé přítelkyni Nikole, jejíž podpora a přínosné podněty mi v práci velmi pomohly.

UNIVERZITA PALACKÉHO V OLOMOUCI

Přírodovědecká fakulta

Akademický rok: 2019/2020

ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: Bc. Martin SADÍLEK
Osobní číslo: R190016
Studijní program: N1301 Geografie
Studijní obor: Geoinformatika
Téma práce: Vymezení výtvarného stylu tematických map metodami machine learning
Zadávací katedra: Katedra geoinformatiky

Zásady pro vypracování

Cílem diplomové práce je identifikovat výtvarné styly několika sad tematických map pomocí programu ORANGE využívající techniky machine learning. Student připraví alespoň tři sady tematických map (geologické, předpovědi počasí, turistické aj.) a zpracuje je především modulem Image Analyst programu ORANGE. Z dosažených výsledků z výtvarného srovnání provede interpretaci z pohledu výtvarného stylu. Následně student všechny mapy kartograficky popíše, provede predikci jejich klasifikace a interpretuje výsledky z kartografického pohledu. Shromážděná data a všechny mapy student přiloží k práci v digitální formě. Dále vyplní údaje o všech datových sadách, které vytvořil nebo získal při řešení cílů práce, do Metainformačního systému katedry geoinformatiky a současně zálohu údajů ve formě validovaného XML souboru. O diplomové práci student vytvoří webovou stránku v souladu s pravidly dostupnými na stránkách katedry a poster o diplomové práci ve formátu A2. Na závěr textu práce připojí student jednostránkové resumé v anglickém jazyce. Celou práci (text, přílohy, poster, výstupy, zdrojová a vytvořená data, XML soubor) odevzdá v digitální podobě na paměťovém médiu a text práce s vybranými přílohami ve dvou svázaných výtiscích na sekretariát katedry.

Rozsah pracovní zprávy: max. 50 stran
Rozsah grafických prací: dle potřeby
Forma zpracování diplomové práce: tištěná

Seznam doporučené literatury:

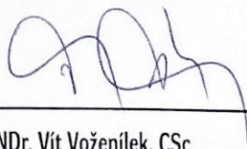
VOŽENÍLEK, V., KAŇOK, J., a kol. (2011): Metody tematické kartografie – Vizualizace prostorových jevů. Univerzita Palackého v Olomouci, 216 s.
DEMŠAR, J. et al. (2013): Orange: data mining toolbox in Python. Journal of Machine Learning Research, 14 (1), s. 2349?2353.
TOPLAK, T. et al. (2017): Infrared Orange: Connecting Hyperspectral Data with Machine Learning, Synchrotron Radiation News, 30(4), s. 40-45.
<https://orange.biolab.si/>
kvalifikační práce studentů katedry geoinformatiky PFF UP z oblasti tematické kartografie a data mining

Vedoucí diplomové práce: prof. RNDr. Vít Voženílek, CSc.
Katedra geoinformatiky

Datum zadání diplomové práce: 7. října 2019
Termín odevzdání diplomové práce: 6. května 2021

L.S.

doc. RNDr. Martin Kubala, Ph.D.
děkan



prof. RNDr. Vít Voženílek, CSc.
vedoucí katedry

OBSAH

SEZNAM POUŽITÝCH ZKRATEK	9
ÚVOD	10
1 CÍLE	11
2 METODY A POSTUP ZPRACOVÁNÍ	12
2.1 Použité metody	12
2.2 Data	13
2.3 Software	15
2.4 Postup zpracování	15
2.5 Zájmová území	17
3 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY	19
3.1 Výtvarný styl	19
3.2 Machine learning	21
3.2.1 Typy algoritmů strojového učení	22
3.2.2 Základní typy úloh v oblasti strojového učení	23
3.2.3 Neuronové sítě	24
3.2.4 Využití neuronových sítí	26
3.3 Orange	28
Hlavní alternativy k softwaru Orange	30
4 KARTOGRAFICKÝ POPIS, PREDIKCE KLASIFIKACE	31
4.1 Dataset FORECAST	33
4.2 Dataset TURIST	33
4.3 Dataset TEKTO	34
5 TVORBA WORKFLOW V ORANGE	38
6 ANALÝZA VÝTVARNÉHO STYLU MAP PŘEDPOVĚDI POČASÍ	40
6.1 Ořez a otočení vybraných map	42
6.2 Barevnost map	43
6.3 Rozlišení map	45
6.4 Výsledky experimentů	48
6.5 Využití nehierarchického shlukování	49
7 ANALÝZA VÝTVARNÉHO STYLU TURISTICKÝCH MAP	51
7.1 Nehierarchické shlukování datasetu TURIST	55
7.2 Výsledky experimentů	57
8 ANALÝZA VÝTVARNÉHO STYLU TEKTONICKÝCH MAP	58
8.1 Nehierarchické shlukování datasetu TEKTO	60
9 VYUŽITÍ VLASTNÍ NEURONOVÉ SÍTĚ V ORANGE	61
10 UŽIVATELSKÉ HODNOCENÍ VÝTVARNÉHO STYLU MAP	63
10.1 Výsledky dotazníkového šetření	66
11 VÝSLEDKY	71
12 DISKUZE	74
13 ZÁVĚR	76
SUMMARY	78
PŘÍLOHY	

SEZNAM POUŽITÝCH ZKRATEK

Zkratka	Význam
AI	Artificial intelligence (umělá inteligence)
ANS	Artificiální neuronové sítě
ANN	Artificial neural network
ANOVA	Analysis of Variance
API	Application Programming Interface (rozhraní pro programování aplikací)
CNN	Convolution neural network
CNTK	Microsoft Cognitive Toolkit
DPI	dots per inch (obrazové body na palec)
GMLMT	Graphic Map Load Measuring Tool
HS	hierarchické shlukování
KNS	Konvoluční neuronové sítě
SNNS	Stuttgart Neural Network Simulator

ÚVOD

Je překvapivé, že v roce 2020 neexistovala žádná podobná studie, která by zkoumala vymezení výtvarného stylu obrazových dat, do kterých mapy spadají, pomocí neuronových sítí, tedy jednou z metod machine learning. Důvodem této situace je pravděpodobně opomenutí machine learning v kartografickém výzkumu, přestože v jiných odvětvích jsou tyto metody hojně využívány. Naděje byly vkládány do programu Orange, protože se již osvědčil při analýze podobnosti evropských měst z Urban Atlasu (Dobešová, 2019).

Práce se zabývá výtvarnými styly tematických map. Pojem výtvarný styl je možné definovat jako provedení grafiky, výběr barev, znakového klíče a provedení layoutu (Vondráková, 2020). Výtvarný styl je tedy jedna z prvních věcí, která je uživatelem vnímána na každé mapě. Problémem je, že neexistuje konkrétní vymezení výtvarných stylů na základě výše zmíněné definice. Metody machine learning by mohly tuto díru zaplnit.

Pro machine learning existuje celá řada softwarových řešení, která umožňují zpracování obrazových dat s využitím neuronových sítí, např. knihovny TensorFlow, Keras či Torch. Software Orange disponuje oproti těmto knihovnám uživatelsky přívětivým prostředím a uživatel není odkázán na své dovednosti v Pythonu. Orange se nabízí pro analýzu obrazových dat metodami machine learning především pro uživatele s menšími zkušenostmi s pokročilým programováním.

Tato diplomová práce je v mnoha ohledech průlomová a poskytuje náhled na potenciál metod machine learning (konkrétně neuronových sítí) a využití softwaru Orange k ještě nerealizovaným kartografickým úlohám.

1 CÍLE

Cílem diplomové práce je identifikovat výtvarné styly několika sad tematických map pomocí programu Orange využívající techniky machine learning. Řešení hlavního cíle je rozděleno do čtyř dílčích postupných cílů.

Prvním dílčím cílem je z dostupných zdrojů získat několik sad tematických map odlišného zaměření, které budou sloužit k analýze jejich výtvarného stylu. Byly zvoleny sady turistických map (TURIST), map předpovědi počasí (FORECAST) a tektonických map (TEKTO). Důvodem pro výběr těchto datasetů byla jejich různorodost, a to především v oblasti jejich barevného a výtvarného provedení. Cílem analýzy bylo zjistit o jednotlivých mapách jejich grafickou náplň, barevné provedení a rozložení základních kompozičních prvků.

Druhým dílčím cílem je kartograficky popsat sestavené sady map a provést experiment „map use“ zaměřený na subjektivní ohodnocení výtvarných stylů map jednotlivých sad. Požadavkem na experiment je jeho nasazení na skupinu vybraných osob a porovnat výsledky subjektivního hodnocení a hodnocení výtvarného stylu tematických map pomocí neuronových sítí ze softwaru Orange.

Třetím dílčím cílem je pomocí softwaru Orange zjistit, jak se mapy seskupí podle výtvarných stylů s využitím již integrovaných neuronových sítí.

Čtvrtým cílem je porovnat výsledky z druhého a třetího dílčího cíle, tj. ohodnocení map uživateli v experimentu vůči výsledkům vyhodnocení map ze softwaru Orange využívajícího neuronové sítě. Porovnáváno bude zařazení jednotlivých map podle výtvarného stylu a také úspěšnost jednotlivých neuronových sítí defaultních pro software Orange. Snahou je sestavit doporučení pro tvorbu neuronové sítě k hodnocení výtvarného stylu nejen tematických map.

2 METODY A POSTUP ZPRACOVÁNÍ

2.1 Použité metody

V průběhu zpracování diplomové práce byly použity následující metody:

Metoda sběru informací a dat

Prostřednictvím serverů **Theses** (theses.cz), **Scopus** (scopus.com) a **Web of Science** (webofknowledge.com) a vyhledáváním kvalifikačních prací na webu katedry geoinformatiky byly vyhledány práce s tematikou neuronových sítí, výtvarného stylu a machine learning, které posloužily k bližšímu pochopení problematiky diplomové práce. Jedná se například o práci *The concept and importance of style in cartography* (2009) litevských autorek Beconyte a Viliuviene z univerzity ve Vilniusu, která pojednává o definování stylů v kartografii. Z oblasti machine learning stojí za zmínku články o používání těchto metod v různorodých oborech. Jedná se o využití při klasifikaci galaxií (Kosiba, 2017), detekce nemocí rostlin (Mohanty a kol., 2016) nebo klasifikace úrovně rakoviny kůže (Esteva a kol., 2017). Dále byly také využity knižní zdroje, a to především v online podobě z důvodu dlouhého uzavření knihoven v důsledku pandemie. Využito bylo například Národní digitální knihovny (ndk.cz). Vyhledávány byly také informace o funkcionalitě softwaru Orange na oficiálním webu tohoto softwaru. Využity byly také další webové stránky, které umožnily získání dalších důležitých informací pro tuto práci, a to především v oblasti neuronových sítí a machine learning. Data pro práci byla stahována z různých webových stránek, jako je například windy.com, yr.no nebo portal.chmi.cz.

Analýza kartografických děl („heuristická analýza“)

Pro kartografický popis mapových děl a následnou potenciální predikci klasifikace dalších map bylo vycházeno z heuristické analýzy. Tato metoda je základní metodou zkoumání použitelnosti webových stránek, kdy se zjišťuje, zda každý prvek uživatelského rozhraní následuje seznam uznávaných principů použitelnosti (heuristik) (Štrupl, 2008). Nejprve byly vybrány nejdůležitější komponenty (heuristiky) mapových děl a ty byly rozděleny do logických hodnotících kategorií (viz kapitola 4). Poté byla ke každé komponentě doplněna hodnota koeficientu (upravená Likertova stupnice) podle toho, zda tato komponenta je součástí dané mapy. Dalším krokem bylo sečtení hodnot komponent u jednotlivých logických kategorií, které by mohlo sloužit v případě kvantifikace barevnosti k tvorbě shluků.

Metody machine learning

V softwaru Orange byly využity již existující nástroje tohoto programu využívající neuronové sítě, pomocí kterých byly analyzovány mapy tří tematických sad, a to mapy turistické (TURIST), předpovědi počasí (FORECAST) a tektonické (TEKTO). Tyto metody jsou definovány jako soubor metod, které umožňují automaticky detekovat vzory v datech a následně předpovídat budoucí data nebo provádět jiné druhy rozhodnutí (Murphy, 2012). Provedené analýzy měly za cíl zjistit schopnost využití těchto nástrojů pro definování výtvarného stylu map. Dále bylo zjišťováno, podle kterých parametrů tohoto stylu dochází ke shlukování, a které parametry vstupních dat mohou být pro tento proces problematické a mohou shlukování negativně ovlivňovat.

Metoda experimentu

Uživatelský experiment byl proveden s využitím metody dotazníkového šetření. Jedná se o kvantitativní metodu výzkumu veřejného mínění, která je hojně využívána například v sociologii (Maříková, 1996). Uživatelům byl rozeslán dotazník s testovacím souborem tematických map. Úkolem těchto uživatelů bylo vytvořit shluky map podle jejich výtvarného stylu na základě jejich vlastního uvážení. Takto vytvořené shluky byly použity pro následné porovnání se shluky vytvořenými neuronovými sítěmi.

Metoda matematické statistiky

Pro vyhodnocení vlivu grafické náplně map na shlukování v prostředí Orange bylo pracováno s metodami matematické statistiky, konkrétně s metodou ANOVA a s Conover-Iman testem. Matematická statistika je dílčí statistická disciplína, která je založena především na teorii pravděpodobnosti (Kladivo, 2013). Využito bylo softwaru Microsoft Excel a RStudio.

Metoda komparace (srovnání)

Pro porovnání definovaných shluků uživateli s výsledky shlukové analýzy byla využita metoda komparace. Shluky vygenerované analýzou byly mezi sebou hodnoceny vizuálním porovnáním. Shluky byly následně popsány a interpretovány.

2.2 Data

Pro diplomovou práci byly shromažďovány mapy tří tematik – mapy turistické (TURIST), předpovědi počasí (FORECAST) a tektonické (TEKTO). Tematické mapy jsou podle Voženilka (2004) definovány jako mapy, kde je upřednostněno podrobné

vyjádření jedné nebo několika vlastností prostorových objektů či procesů na úkor vlastností nepodstatných. Tematické mapy s měřítkem od 1 : 5 000 až 1 : 50 000 jsou označovány jako mapy účelové. Pro získávání map byly upřednostněny online zdroje. Důvodem byla snaha získat vstupní data, která by nebyla ovlivněna metodou sběru (např. stíny vzniklé při skenování papírových map v místech přehybů). Dalším důvodem byla také situace kolem světové pandemie a výrazné ztížení získávat mapy mimo internetové prostředí. Získané mapy v Orange vstupují do procesu analyzování výtvarného stylu jako obrázky. Funkcionalita softwaru Orange a vliv tří parametrů vstupních obrazových dat, které by mohly mít vliv na výsledné shlukování vstupních map, byly testovány na mapách předpovědi počasí (konkrétně mapy teploty vzduchu). Výše zmíněnými parametry jsou:

- Rozlišení:
 - Jeden z parametrů, u kterého bylo potřeba zjistit vliv na výsledné zařazení map do shluků, je rozlišení jednotlivých map. Ať už v důsledku rozdílné velikosti vstupních dat nebo v důsledku snížení rozlišení záměrně pro účely testování.
 - Bylo pracováno s hodnotou rozlišení v pixelech.
- Celek vs. výřez mapy:
 - Dalším parametrem, u kterého bylo potřeba ověřit jeho vliv na výsledné shlukování, je případné ořezání originálních dat.
- Natočení map:
 - Natočení mapy je parametr, který by mohl odhalit problémy u vstupních dat například z důvodu využití rozdílných zobrazení.

Dále byla na mapách předpovědi počasí testována jejich barevnost jako jeden z prvků výtvarného stylu:

- Barevný prostor:
 - Jeden z hlavních parametrů při hodnocení výtvarného stylu je barevné provedení map. Z tohoto důvodu byl zkoumán barevný prostor map, a to v podobě hledání vlivu dvojic shodných map v barevném a černobílém provedení na výsledné shlukování.

Další dvě sady tematických map byly využity pro testování dalších parametrů (grafická náplň map) ovlivňující shlukování, pro zjištění vlivu vybrané shlukovací metody a také pro ověření nebo vyvrácení výsledků z analýzy map předpovědi počasí. Nakonec byly všechny sady použity pro kartografický popis, uživatelský experiment

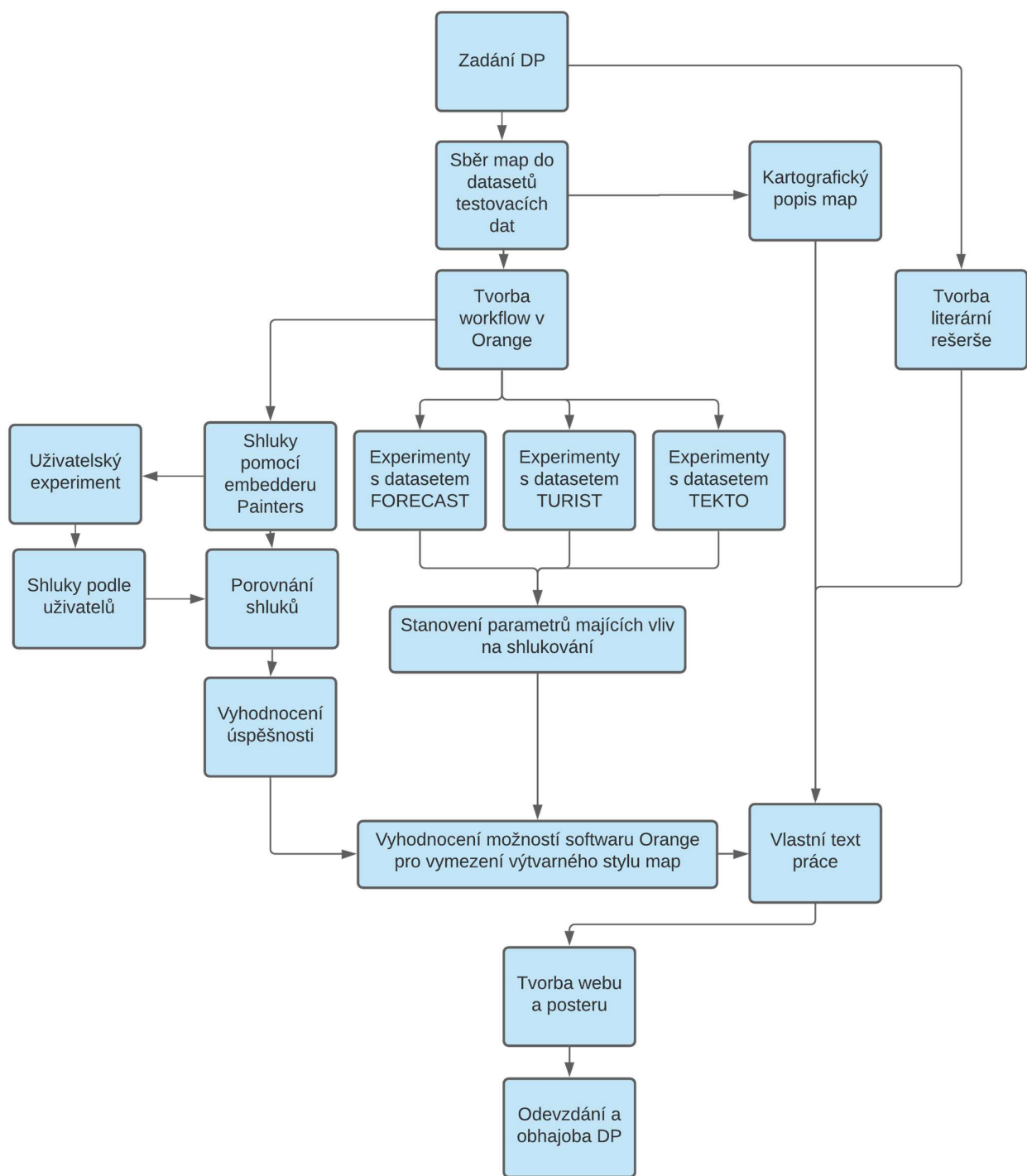
a k tvorbě výsledných shluků pro porovnání a vyhodnocení využití pro definování výtvarných stylů map.

2.3 Software

Hlavní část práce byla provedena v softwaru **Orange**. První práce na vývoji softwaru umožňujícím pracovat s metodami strojového učení provedl ve Slovinsku institut Jožefa Stefana ve spolupráci s Univerzitou Ljublaň v roce 1996. Práce sice z počátku probíhaly v prostředí jazyka C++, ale první verze frameworku byla postavená na Pythonu a byla představena 10. října 1996. V roce 1997 byl poté vytvářený framework na C++ spojen s moduly z Pythonu a od tohoto roku nese software název Orange. V roce 2005 bylo přidáno graficky přívětivé uživatelské prostředí a v roce 2005 extenze pro datovou analýzu a bioinformatiku. V roce 2009 byl vydán Orange 2.0. Méně zkušené uživatelské přivítali v roce 2013 velké změny uživatelského prostředí. V současné době (březen 2021) je k dispozici Orange ve verzi 3.28.0 (URL9). Pro práci s tabelárními daty ze softwaru Orange, bylo pro základní statistické vyhodnocení dotazníkového šetření a pro analýzu dat pomocí statistické metody ANOVA využito softwaru **Microsoft Excel**. Dále bylo využito softwaru **Zoner Photo Studio 14** pro úpravu map do vhodné podoby (změna rozlišení, barevnost) pro experimenty v softwaru Orange. Pro náročnější analýzy matematické statistiky bylo využito softwaru **RStudio** ve verzi 1.4 pracující s **R** ve verzi 4.0.5. Zjištění grafické náplně map bylo provedeno nástrojem *Graphic Map Load Measuring Tool* v softwaru **GIMP**.

2.4 Postup zpracování

Prvním krokem po zadání diplomové práce byl sběr map do datasetů a literární rešerše problematiky. Mapy všech datasetů (FORECAST, TURIST a TEKTO) byly následně kartograficky popsány. Dále bylo v softwaru Orange vytvořeno workflow pro zpracování dat pomocí metod machine learning s využitím modulu Image Analytics. Pomocí něho byly mapy z testovacích datasetů analyzovány pomocí dostupných embedderů, ze kterých byl vybrán embedder Painters. Na jednotlivých datasetech byla následně provedena série experimentů, které měly za cíl zjistit, jaký je vliv parametrů výtvarného stylu a parametrů map na proces embeddingu a shlukování. Mapy datasetů také vstupovaly do uživatelského experimentu, z něhož vytvořené shluky byly následně porovnány se shluky vytvořených pomocí embedderu Painters. Na závěr byly vyhodnoceny možnosti softwaru Orange pro vymezení výtvarného stylu map. Postup a výsledky byly poté sepsány do textu práce a byl vytvořen web a poster pojednávající o diplomové práci.



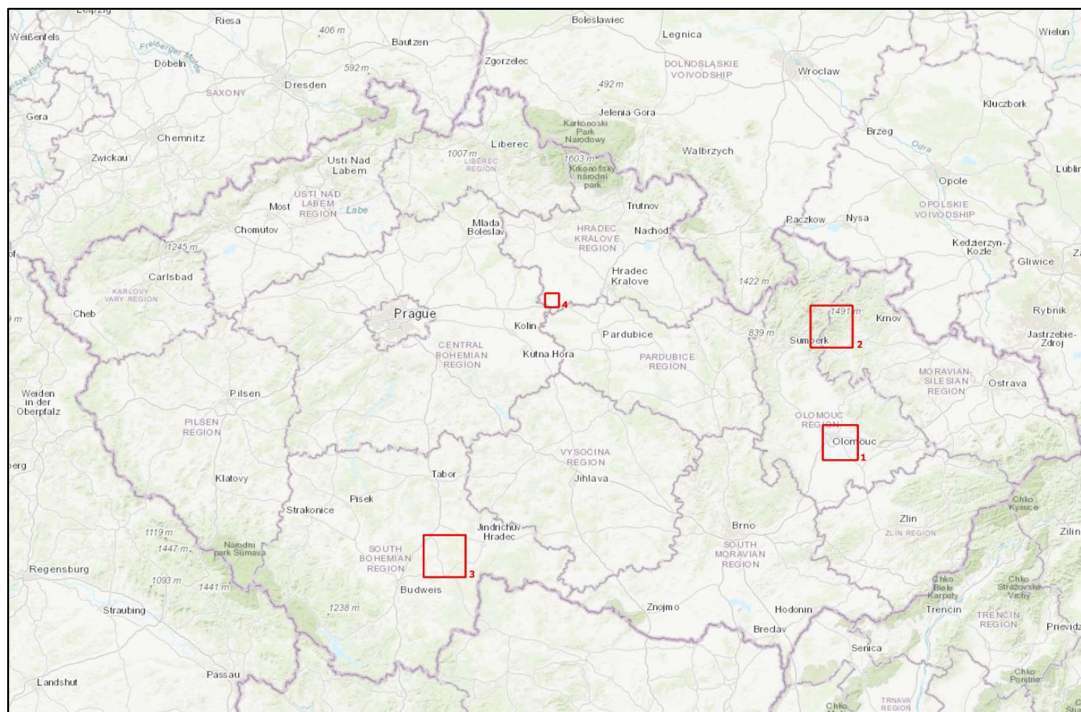
Obr. 2.1 Schéma postupu zpracování diplomové práce.

2.5 Zájmová území

Pro každý dataset bylo vybráno jiné zájmové území. Pro hodnocení jednotlivých turistických map z datové sady TURIST byla zvolena čtyři území s odlišnou charakteristikou krajiny. Z těchto území byl hodnocen výtvarný styl turistických map (tab. 1, obr. 2.2). Hlavním kritériem výběru území byla odlišnost jejich krajiny z pohledu funkčního využití/vzhledu/odlišnosti kartografického obrazu. Mapy byly získávány v lednu 2021.

Tab. 1 Vybraná zájmová území pro sadu map TURIST.

Označení území	Území	Charakteristika krajiny
1	Olomouc a blízké okolí	městská zástavba
2	Praděd a okolí	horská oblast
3	Jihočeské rybníky	rybníkářská oblast
4	Chlumeck nad Cidlinou a okolí	zemědělská nížinná krajina



Obr. 2.2 Vybraná území pro sadu map TURIST.



Obr. 2.3 Ukázka odlišnosti vybraných území na online mapách.

Pro hodnocení výtvarného stylu sady map předpovědi počasí FORECAST bylo zvoleno území Evropy. Staženy byly mapy teploty vzduchu v odpoledních a brzkých ranních hodinách v druhé polovině září 2020. Pro další analýzu se pracovalo s deriváty originálních map, kdy byly vyříznuty dílčí části území, zmenšeny originální data a také byla vybraná data otočena.

Pro třetí skupinu tematických map byly vybrány mapy tektonické (TEKTO), konkrétně mapy tektonických desek světa. Zobrazovaným územím je tedy celý svět. Díky tomu byla pro každou datovou sadu využita jiná oblast a byla získána co největší různorodost dat. Časová složka vzniku těchto map nehraje roli, tektonické desky se s měnícím časem pohybují velmi pomalu a na všech mapách je současný stav světové tektoniky.

3 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY

3.1 Výtvarný styl

Podle ústního sdělení doktorky Aleny Vondrákové (ze dne 2. března 2020) je nutné rozlišovat výtvarný (grafický) styl, umělecký styl a kartografický styl map. **Výtvarný** styl je provedení grafiky (ta může být například minimalistická), dále provedení kompozičních prvků a jejich parametrů (ukončení linií, jejich tloušťky, výběr fontu...), výběr barev, znakového klíče a provedení layoutu (výtvarný styl je také možné rozdělit na styl mapy a styl kompozice). **Umělecký** styl odpovídá teorii výtvarného umění (např. impresionismus, romantismus). Týká se spíše provedení celého díla nežli jeho dílčích částí. Pravda ho charakterizuje „nadřazením grafických až malířských prvků v úloze stylistických prostředků“ (Pravda, 2003, s. 90). V Tezauru je umělecký styl definován jako „styl konkrétního umělce, školy nebo hnutí“ (Visual Thesaurus, 2021). Posledním je **kartografický** styl, který ve své práci dále více rozebírala Drahošová (2015). Ta zde uvádí několik definic, mezi nimi i definici Voženilka, Kaňoka a kol. (2011), že se jedná o ustálený způsob výběru některých kartografických metod zpracovatelem mapy. Ten je však často ovlivňován jak objektivními, tak i subjektivními stylovými faktory.

V bakalářské práci Bači (2009) je výtvarný styl map rozdělen na styl **umělecký** a styl **populární**. Bača čerpal z knihy *Mapový jazyk* od Pravdy (1997), který zde rozděluje umělecký styl na švýcarskou školu a na panoramatický styl, které jsou ovlivněny hlavně E. Imhofem, a na malířský umělecký styl.

Uměleckým stylem map se dá také nazvat forma uměleckého vyjádření, které popisují ve své práci Ribeira a Caquarda (2018, s. 1) jako: „umělecké projevy, které využívají mapy a kartografické prvky estetickým způsobem, většinou pro politické účely, s využitím několika technik jako je malba, sochařství, rytina, fotografie, koláž, kresba, performance nebo instalace“. V tomto případě je mapa využita pouze jako sekundární prvek a jedná se pouze o prostředek vizuálního efektu.

Kromě výzkumu stylů map v České republice byl prováděn výzkum také v zahraničí, a to na univerzitě ve švýcarském Curychu. Ory, Christophe a Fabrikant se ve své práci s názvem *Identification of styles in topographic maps* (2013) snažili identifikovat styly topografických map. Ve své práci pracují s francouzskými a švýcarskými topografickými mapami, na kterých se snaží popsat **topografické** styly těchto map pomocí metody vizuálního srovnání. Z práce vyplývá, že jako topografický styl je považován **kartografický** styl mapy. Pro toto srovnání použili mapy stejné

oblasti od dvou různých vydavatelů, a to od francouzského IGN a švýcarského Swisstopo. Při porovnání obou map se zaměřili především na rozdíly v barevném provedení map a na rozdíly v typografii. Výsledkem práce jsou parametry, které odlišují kartografické styly těchto map. Jedná se například o barevné provedení, rozdílné použití bodových prvků nebo využití stínovaného reliéfu.

Další prací zabývající se styly map je *The concept and importance of style in cartography* (2009) od autorek Beconyte a Viliuviene z univerzity ve Vilniusu. Ve své práci definují na základě analýzy kartografických znaků čtyři hlavní skupiny stylů map, a to:

- minimalistické,
- standartní,
- konvenční,
- nápadné.

Hodnotícími parametry pro předchozí rozdělení jsou dekorativnost, expresivita, originalita a konzistentnost map. Podle těchto parametrů provedli také následující rozdělení na styly:

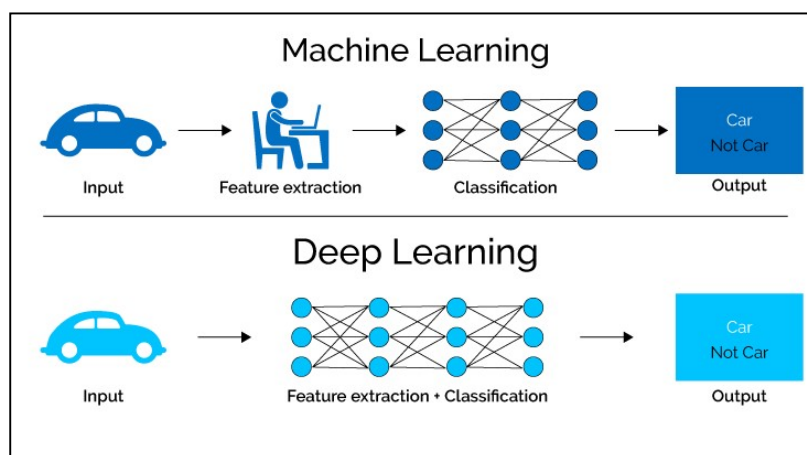
- konvenční (conventional),
 - lakonický (laconic),
 - konstruktivní (constructive),
- expresivní (expressive),
 - expresivní (expressive),
 - lehký (lightsome),
 - agresivní (aggressive),
- umělecké (artistic),
 - starobylý (antique),
 - měkký/lyrický (soft/lyric),
 - lahodný (luscious),
 - extravagantní (extravagant).

Pro rozdělení do podkategorií předchozích třech stylů je nutné využít další parametry map jako jsou například kontrast nebo historické a kulturní asociace. Výše uvedené styly jsou zastřešeny termínem **kartografický styl**. Podle definic uvedených na začátku kapitoly by bylo nicméně vhodnější zastřešit tyto pojmy pod **výtvarný styl**. Využití různých stylů není příliš patrné u topografických map, větší různorodost můžeme nalézt u map tematických.

Z výše uvedeného je patrné, že přesná definice výtvarného stylu v kartografii je poměrně nejednoznačná. Problematický je také překlad tohoto termínu do angličtiny. Neexistuje totiž termín pro výtvarný styl, ale pouze pro umělecký styl. Východiskem pro práci je definice výtvarného stylu Vondrákové uvedené výše.

3.2 Machine learning

Machine learning, česky strojové učení, je oblast matematické informatiky, konkrétněji umělé inteligence (anglicky artificial intelligence – AI), která studuje automatizované metody analýzy dat. Definován je jako soubor metod, které jsou schopné automaticky detekovat vzory v datech a poté předpovídat budoucí data nebo provádět jiné druhy rozhodnutí (Murphy, 2012). Tyto metody umožňují systémům automaticky se učit a zlepšovat své dovednosti ze získaných zkušeností, aniž by byly explicitně naprogramovány (URL1). Strojové učení používá algoritmy vytvořené pro konkrétní úkoly (URL17). Pokud algoritmus AI vrací nepřesnou predikci, musí tvůrce manuálně zasáhnout a provést úpravy (URL18). Podmnožinou machine learning je deep learning (česky hluboké učení). To oproti strojovému učení využívá více vrstev, složených z neuronů, kde každá vrstva používá jako vstup výstup z předchozích vrstev. V deep learning se využívají hluboké neuronové sítě (CNN). Tento přístup více napodobuje činnost lidského mozku (URL17). V tomto případě je algoritmus také sám schopen zjistit, zda je předpověď přesná nebo ne, a to prostřednictvím vlastní neuronové sítě (URL18).



Obr. 3.1 Porovnání strojového učení (nahore) s hlubokým učením (dole). (zdroj: <https://www.xenonstack.com/blog/log-analytics-deep-machine-learning/>, 2021)

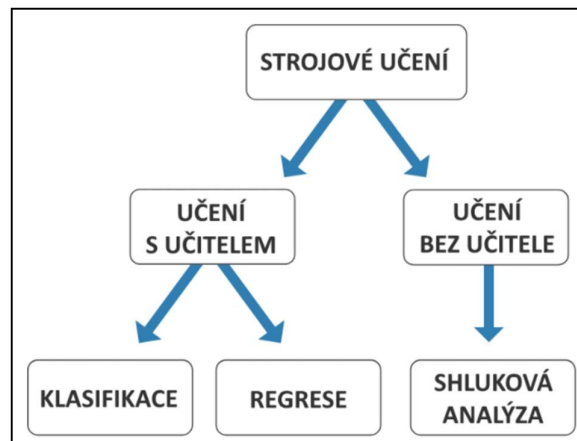
Machine learning se prolíná s mnoha oblastmi, mezi ně patří i data mining (česky dolování dat). Často používají stejné metody a tyto dvě oblasti se významně překrývají, ale zatímco machine learning se zaměřuje na predikci na základě

známých vlastností získaných z tréninkových dat, data mining Berka (2003) popisuje jako: „*netriviální získávání implicitních, dříve **neznámých** a potenciálně užitečných informací z dat*“.

3.2.1 Typy algoritmů strojového učení

V rámci strojového učení existuje několik typů algoritmů (Mohri et al., 2012):

- učení s učitelem – stroj pracuje se sadou trénovacích dat obsahující dvojici - data a popisky; popisky chybí v testovacích datech a algoritmus se je na základě trénovacích dat snaží predikovat; typickým zástupcem učení s učitelem je shluková analýza,
- učení bez učitele – stroj pracuje s trénovacími daty bez popisků; predikce se vytváří pouze na základě nepopsaných dat; náročnější na vyhodnocení efektivity učení; do této kategorie spadají klasifikace a regrese,
- kombinace učení s učitelem a bez učitele – využívají se trénovací data s popiskami i bez popisků,
- další typy učení, např. zpětnovazebné učení, online učení, transduktivní učení nebo aktivní učení.



Obr. 3.2 Základní typy úloh v oblasti strojového učení a využívané typy učení (zdroj: <https://www.humusoft.cz/blog/20180817-strojove-uceni/>).

3.2.2 Základní typy úloh v oblasti strojového učení

Výše zmíněné typy učení jsou schopné různých typů úloh (obr. 3.2), Mezi základní typy úloh v oblasti strojového učení patří (URL19):

Klasifikace

Při klasifikaci objektů je zadáno klasifikační kritérium nebo pravidla pro rozklad objektů do kategorií (Šarmanová, 2012). V těchto úlohách se využívají trénovací data, která obsahují vstupní atributy. Ty jsou ve struktuře řádků a sloupců spolu s vektorem cílových atributů vyjadřujícími informaci o zařazení objektů do tříd nebo predikovanou hodnotu. Výsledky jsou využitelné při klasifikaci nových objektů přiřazením hodnot vstupních atributů k cílovému atributu. Pokud je klasifikace založena na rozhodovacích stromech, je výstupem klasifikace právě rozhodovací strom. Pokud je rozhodnutí o klasifikaci založené na neuronové síti, znázorněno je vahami vazeb mezi neurony (Berka, 2003).

Regrese

Cílem regresní analýzy je odhad spojitě výstupní veličiny na základě vstupních dat (URL19). Na základě vstupních trénovacích dat dochází k hledání matematické funkce, která by dokázala co nejpřesněji vyjádřit křivku procházející těmito daty. Aproximací hledané funkce se nejčastěji dosahuje metodou nejmenších čtverců (Berka, 2003).

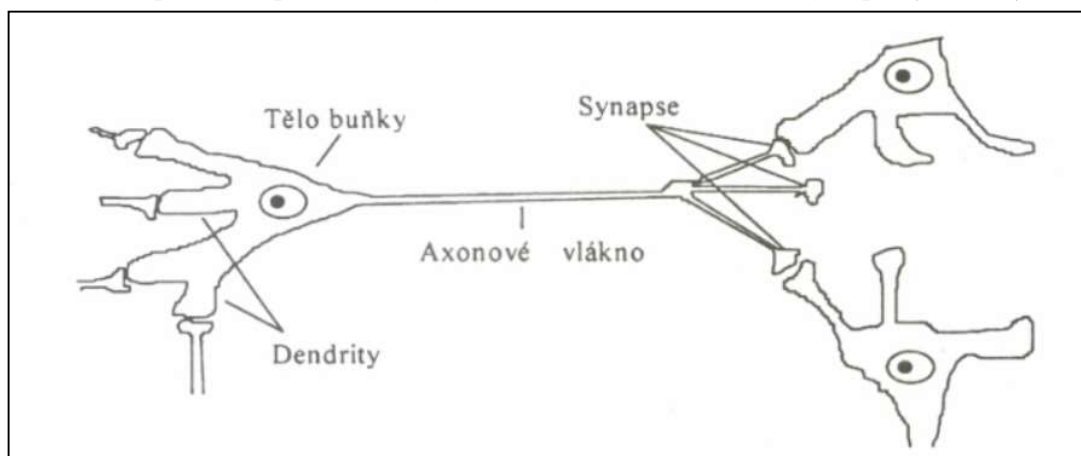
Shluková analýza

Při shlukové analýze nejsou kritéria shlukování známa nebo jsou předmětem výzkumu (Šarmanová, 2012). Cílem je shlukování objektů do podobných skupin. Proces shlukování se provádí pomocí řady různých statistických algoritmů a metod. Shluková analýza nerozlišuje mezi závislými a nezávislými proměnnými. Analýzou všech objektů a jejich vzájemných vztahů se snaží objevit podobnostní vztahy s cílem identifikovat jednotlivé shluky. Mezi dvě základní metody shlukování patří **hierarchické** a **nehierarchické** shlukování. První zmíněná metoda generuje shluky, které jsou organizovány do hierarchické struktury. Tato technika vyžaduje pouze míru podobnosti mezi objekty. Nenastavuje se množství shluků. Tuto metodu je dále možné rozdělit na hierarchické shlukování **aglomerativní** a **divizní**. Aglomerativní shlukování je bottom-up přístup. Každý objekt je na začátku shlukování považován za samostatný shluk a poté dochází ke kombinaci objektů do větších shluků na každé úrovni podle jejich podobnosti. Divizní shlukování využívá naopak top-down přístupu.

Na začátku uvažuje všechny objekty jako jeden velký shluk a až poté dochází k rozdělení tohoto shluku do menších shluků na každé úrovni. Při použití některé z metod nehierarchického shlukování nedochází k vytváření hierarchické struktury, ale rozkládá se pouze daná množina objektů do podmnožin dle předem daného kritéria. Příkladem nehierarchického shlukování je metoda K-means (Dean, 2014).

3.2.3 Neuronové sítě

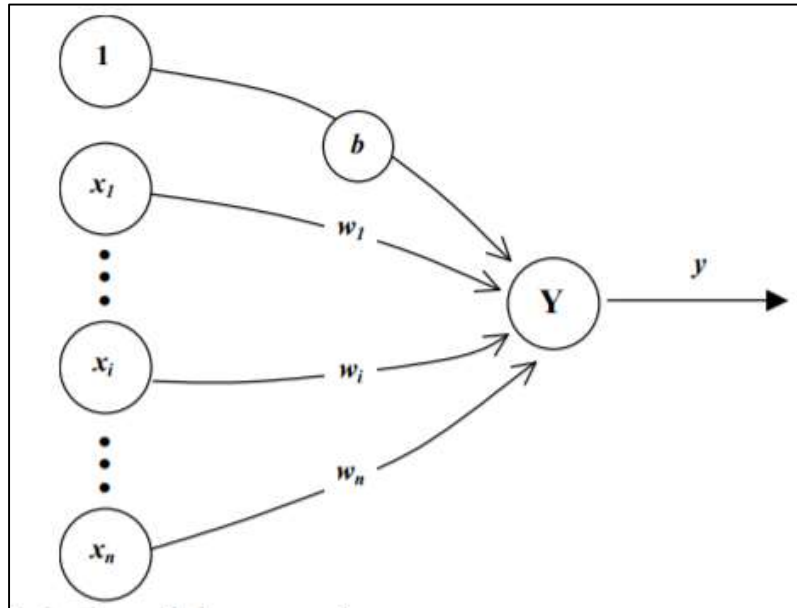
Jednou z metod umělé inteligence jsou neuronové sítě. Za teoretický základ o neuronových sítích posloužil učební text *Neuronové sítě 1* (Volná, 2008). Neuronové sítě jsou jednou ze základních technik strojového učení, které umožňují systémům učit se z vkládaných dat. Algoritmy neuronových sítí jsou inspirovány způsobem zpracování informací v biologickém nervovém systému. V lidského mozku jsou mezi sebou propojeny jednotlivé neurony s dendrity (vstupní větve, ze kterých přicházejí informace), které jsou schopné přijímat, vést a zpracovávat podněty. Z těchto vstupů poté produkují výstup přes axon do jiného neuronu. Propojení s dalšími neurony je umožněno pomocí speciálních výběžků na dendritech, tzv. synapsí (obr 3.3).



Obr. 3.3 Biologický neuron (Volná, 1998).

Neuronová síť je tvořena tzv. formálními neurony. Jedná se o umělé neurony tvořené z n reálných vstupů označovaných jako x . Vstupy jsou ohodnoceny pomocí synaptických vah, tvořících vektor $w = (w_1, \dots, w_n)$. Dalším parametrem je bias neboli práh, označovaný písmenem b . Ten označuje hodnotu, při které dojde k aktivaci neuronu aktivací funkcí. Nejčastěji používanou je sigmodiální funkce. Při aktivaci dojde k přiřazení hodnoty výstupu y . Na základě správnosti výstupu y se upravuje hodnota vah. Pokud je výstup správný, váhy se nemění. Pokud je výstup 0 a předpoklad (d) je roven 1, dojde k posílení vah. V opačném případě dojde k oslabení vah. Neuronové sítě jsou uspořádány z formálních neuronů. Ty jsou mezi sebou navzájem spojeny a výstup jednoho neuronu je vstupem do dalších neuronů.

Správně by měla být navržena jedinečná neuronová síť pro řešení konkrétní úlohy. Tato práce tedy zjišťuje, zda je možné již existující neuronovou síť efektivně použít za jiným účelem, než za kterým byla původně vytvořena.



Obr. 3.4 Model formálního neuronu s biasem (zdroj: Volná, 1998)

Neuronové sítě lze rozdělit na dva typy a to na sítě konvoluční (KNS) - CNN (Convolutional Neural Network) a umělé - umělé (ANS) - ANN (Artificial Neural Network). KNS jsou určeny pro zpracování dat s mřížkovitým vzorem (Yamashita, 2018). Tyto sítě nacházejí nejčastěji uplatnění při rozpoznávání obrazových dat (O'Shea, Nash, 2015). Krogh (2008) popisuje ANN jako jednu z metod machine learning, která byla vytvořena na základě lidského mozku, ve kterém probíhá komunikace pomocí neuronů, a to díky vysílání elektrických impulsů skrz síť tvořenou z těchto neuronů. Hlavní rozdíl mezi KNS a ANN spočívá v tom, že pouze poslední vrstva u KNS je plně spojena, zatímco u ANN každý neuron v jedné vrstvě plně spojen s každým neuronem v další vrstvě. (Gogul, Kumar, 2017). ANS se využívají především na omezené datové sady a na neobrazová data. KNS jsou výkonnější a poskytují přesnější řešení klasifikačních problémů u obrazových dat. (URL13). Software Orange pracuje v rámci nástroje Neural Network s ANS, konkrétně s vícevrstevným perceptronem (URL14). Nástroj Image Embedding naopak pracuje s KNS (URL11).

3.2.4 Využití neuronových sítí

Využití neuronových sítí a metod machine learning je velice široké. Uplatnění lze nalézt v mnoha různorodých oblastech. Pro příklad rozsáhlosti zastoupených témat v rámci analýzy obrazových dat stojí za zmínku klasifikace galaxií (Kosiba, 2017), detekce nemocí rostlin (Mohanty a kol., 2016) nebo klasifikace úrovně rakoviny kůže (Esteva, 2017). Oblast klasifikace map pomocí metod machine learning ale bohužel není příliš rozpracovaná.

Studie *Deep Convolutional Neural Networks for Map-Type Classification* autorů Zhou, Li, Arundel a Liu z Cornellské univerzity z roku 2018 popisuje využití konvoluční neuronové sítě pro velmi rychlou klasifikaci typu map. Pro svůj výzkum použili sedm různých kategorií map (topografické mapy, 3D mapy, ortofoto snímky...) z online zdrojů, např. ArcGIS a Google Maps. Každá kategorie obsahovala přibližně 200 různých map. Celkový počet map ve studii dosáhl 1812. Výsledkem experimentu byla schopnost neuronové sítě správně zařadit až 70 % map, resp. až 98 % map u hluboké neuronové, a to podle zvoleného typu sítě. Pro zvýšení přesnosti bylo použito větší množství správně popsanych dat pro natrénování neuronové sítě nebo využití pouze hluboké neuronové sítě, která je ale náročná na natrénování.

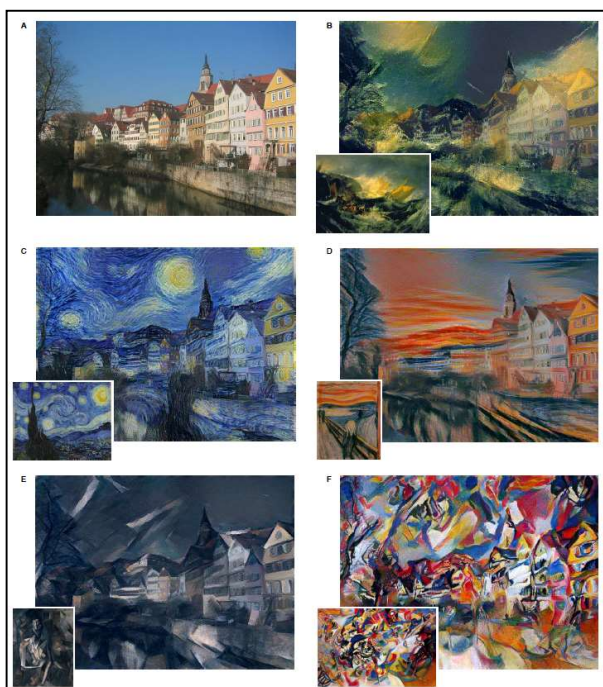
Využití neuronových sítí pro hledání podobnosti evropských měst provedla Dobešová (2019). Vytvořila 100 kruhových výsečí různých evropských měst vyřezáním z Urban Atlasu a analyzovala jejich podobnost v softwaru Orange s využitím nadstavby Image Analytics. Pro analýzu těchto dat vybrala ze všech nabízených embedderů z nástroje *Image embedder Painters embedder*, který byl vytvořen pro určování uměleckých děl malířů a natrénován na 79 433 instancích obrazů od 1 584 malířů (Dobešová, 2019, 23kk). Za vzdálenostní metriku použila kosinovou vzdálenost. V rámci nástroje *Hierarchical clustering* využila Wardovu metodu. Výsledkem bylo rozdělení vstupních dat do čtyř různých skupin podle funkčního členění a struktury jednotlivých měst.



Obr. 3.5 Proces shlukování na uměleckých dílech embedderem Painters (zdroj: <https://github.com/inejc/painters/blob/master/README.md>).

Neuronové sítě využívá i řada online aplikací vytvořených za účelem určování druhu rostlin. Nejznámější je pravděpodobně webová aplikace Pl@ntNet (<https://identify.plantnet.org/>). Aplikace obsahuje přes 22 000 druhů rostlin na více než 3 miliónech snímků. K dispozici je také mobilní verze pro rozhraní Android i iOS. Podobná webová aplikace i s mobilní verzí vznikla na Masarykově univerzitě v Brně. Verze pro web nese název Plant.id (dostupné na: <https://plant.id/>), mobilní aplikace se nazývá FlowerChecker.

Pomocí konvolučních neuronových sítí je také možné provést transfer vizuálního/kartografického stylu (obr. 3.6). Základem pro provedení transferu je extrakce textury ze vstupního obrázku a následná aplikace této struktury na požadovaný obraz (Gatys, a kol., 2016).



Obr. 3.6 Transfer vizuálního stylu z obrazů na vstupující fotografii (zdroj: Gatys, a kol., 2016).

Rozpoznáváním vozidel se ve své diplomové práci (2016) zabýval Jan Sedlák z Masarykovy univerzity v Brně. Zkoumal možnosti využití metod strojového učení při rozpoznání barev a výrobce vozidel z kamerových snímků na území České republiky. Ve své práci použil mělké neuronové sítě.

3.3 Orange

Pro zpracování cíle diplomové práce byl využit open-source software Orange, který je určen především pro data mining a strojové učení. Díky své jednoduchosti a grafickému provedení Orange nachází uplatnění u začátečníků i profesionálů (Demšar a kol., 2013). Nikde se však nenachází statistika počtu stažení či množství uživatelů. Po nainstalování je k dispozici k použití pět základních modulů (*Data*, *Visualize*, *Model*, *Evaluate* a *Unsupervised*), zbylé moduly je nutné doinstalovat. To je případ i modulu *Image Analytics*, který byl primárně použit v této diplomové práci.

Pro analýzu obrazových dat a jejich klasifikaci existuje řada alternativních řešení, např. knihovny TensorFlow, Keras či Torch. Tyto knihovny nedisponují uživatelsky graficky přívětivým prostředím, jako je tomu u Orange, a je nutné s nimi pracovat přímo v rozhraní pro programovací jazyk Python.

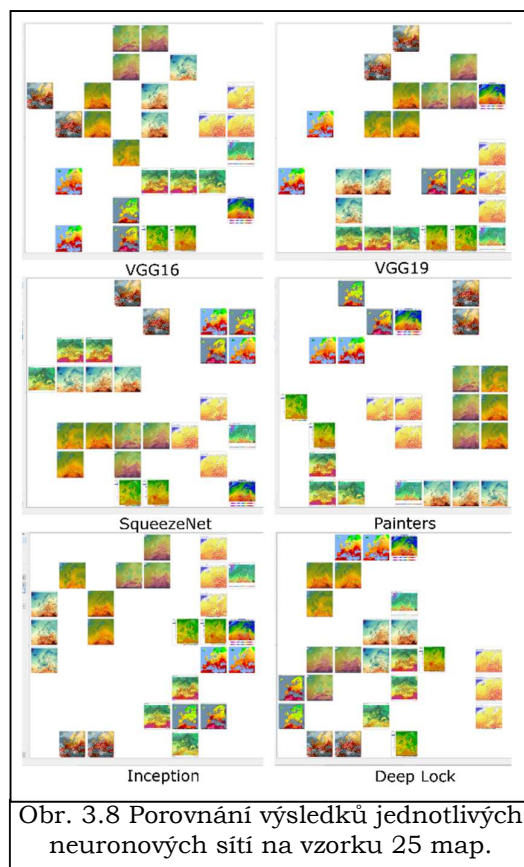
V článku *Democratized image analytics by visual programming through integration of deep models and small-scale machine learning* od Godeca a kol. (2019) o využití neuronových sítí v biomedicině je proces embeddingu popsán jako převod obrazového vstupu na jeho vektorovou reprezentaci. Celkem je při tomto procesu vytvořeno 2 048 vektorů. Součástí extenze *Image Analytics* je nástroj *Image Embedder*, který umožňuje obrazovou analýzu naimportovaných snímků pomocí embeddingu, a to buď lokálně v počítači uživatele, nebo posláním na vzdálený server, kde jsou snímky následně zpracovány. Zástupně lokálně pracujícího embedderu je v softwaru Orange pouze SqueezeNet. Jeho největší výhodou je rychlé zpracování snímků. Pro ostatní embeddery je nutné vzhledem k práci na vzdáleném serveru aktivní internetové připojení. Obrázky poslané na server se nikde neskladují (URL2). Pokud není nastaven lokální embedder, dojde k automatickému přepnutí na embedder SqueezeNet, tedy jediný lokální embedder v Orange. V případě znovuspuštění konkrétního embedderu na stejné snímky dojde ke vzniku stejných shluků.

- *Embedder:*
 - SqueezeNet: **Small and fast** model for image recognition trained on ImageNet.
 - Inception v3: **Google's Inception v3** model trained on ImageNet.
 - VGG-16: **16-layer image recognition model** trained on ImageNet.
 - VGG-19: **19-layer image recognition model** trained on ImageNet.
 - Painters: A model trained to **predict painters from artwork images**.
 - DeepLoc: A model trained to analyze **yeast cell images**.

Obr. 3.7 Výčet defaultních neuronových sítí softwaru Orange (zdroj: University of Ljubljana, 2021).

Při nastavení nástroje Image Embedder je nejprve nutné vybrat atribut, na základě kterého budou data popisována, a poté již preferovaný embedder z nabídky (obr. 3.7). Jejich popis je převzat z oficiální dokumentace softwaru Orange (URL2):

SqueezeNet je hluboký model pro rozpoznávání obrázků, který dosahuje přesnosti na úrovni AlexNet (konvoluční neuronová síť s osmi vrstvami) na datové sadě ImageNet s 50x menším počtem parametrů. Hluboká neuronová síť **Inception v3** společnosti Google je určena k rozpoznávání obrázků. Je naučena na datové sadě ImageNet. **VGG-16** a **VGG-19** jsou hluboké neuronové sítě pro rozpoznávání obrazu od společnosti Visual Geometry Group z University of Oxford. Stejně jako předchozí jsou naučeny na datové sadě ImageNet. **Painters** je embedder, který byl naučen na 79 793 obrazech od 1 584 malířů a vyhrál soutěž Kaggle's Painter by Numbers v roce 2016/2017. Jeho tvůrcem je Nejc Ilenic, softwarový inženýr se zaměřením na machine learning a statistiku ze slovinské Ljublaně (URL15). Konvoluční síť **DeepLock** byla natrénovaná na 21 882 obrazech jednotlivých



Obr. 3.8 Porovnání výsledků jednotlivých neuronových sítí na vzorku 25 map.

buněk, které byly ručně přiřazeny k jednomu z 15 lokalizačních oddílů. V tomto výčtu chybí embedder **openface** pro rozpoznávání obličejů pomocí Python a Torch s hlubokými neuronovými sítěmi (URL10). Tuto síť není nicméně možné účinně použít na mapová díla, protože se na nich nenachází lidský obličej.

ImageNet

ImageNet je obrazová databáze organizovaná podle WordNet (velká lexikální databáze anglických sloves, podstatných jmen, přídavných jmen a příslovců (<https://wordnet.princeton.edu/>)) hierarchie, kde každý uzel této hierarchie je popsán stovkami a tisíci obrázky (URL3). ImageNet využívá pouze podstatná jména a v současné době (duben 2020) se u každého uzlu v průměru nachází přes 500 snímků. Celkem je zde možné najít více než 14 milionů obrázků. Databázi lze stáhnout a používat ji především pro výzkumné a edukativní činnosti. Na tvorbě

databáze ImageNet se podílejí Stanford Vision Lab, Stanford University a Princeton University.

Hlavní alternativy k softwaru Orange

Mezi hlavní alternativy softwaru Orange patří **Keras**. Jedná se o multiplatformní open-source API (Application Programming Interface) pro tvorbu neuronových sítí napsané v Pythonu. Knihovna je pod licenci MIT a je součástí TensorFlow (URL5). Open source platforma **TensorFlow** pro strojové učení obsahuje velké množství nástrojů, knihoven a dalších zdrojů, které využívají nejmodernější technologie v machine learning (URL6). Pro práci s neuronovými sítěmi je možné použít také open-source knihovnu **TyTorch**, která je postavená na knihovně Torch (URL7). Další alternativou je **CNTK** (Microsoft Cognitive Toolkit). Jedná se o open-source sadu nástrojů, která umožňuje uživateli snadno realizovat a kombinovat oblíbené typy modelů, jako jsou dopředné, konvoluční a rekurentní neuronové sítě. S CNTK má podobu knihovny, s kterou je možné pracovat v programech na bázi Python, C# nebo C++ (URL8).

4 KARTOGRAFICKÝ POPIS, PREDIKCE KLASIFIKACE

Pro kartografický popis a následnou predikci klasifikace map jednotlivých datasetů byl využit koncept heuristické analýzy webových stránek, který byl aplikován na mapová díla. Heuristická analýza je základní metodou zkoumání použitelnosti webových stránek, kdy se zjišťuje, zda každý prvek uživatelského rozhraní následuje seznam uznávaných principů použitelnosti (heuristik) (Štrupl, 2008). Nejprve byly vybrány nejdůležitější prvky (heuristiky) mapových děl jednotlivých tematik. Pokud některý z prvků byl součástí jen vybraných tematických map, je tato skutečnost zaznačena pomocí zkratk u těchto prvků. Prvky byly rozděleny do čtyř logických hodnotících kategorií, a to základní a nadstavbové kompoziční prvky podle Voženilka (2004), na grafické provedení mapy a další prvky, které se na mapách vyskytují:

- **Základní kompoziční prvky:**
 - název (titul/podtitul),
 - legenda,
 - měřítko (číselné/grafické),
 - tiráž,
 - mapové pole.
- **Nadstavbové kompoziční prvky:**
 - směrovka,
 - logo,
 - tabulky,
 - grafy,
 - vedlejší mapy,
 - obrázky,
 - textová pole,
 - blokdiagramy,
 - citace,
 - rejstříky a seznamy,
 - reklamy.
- **Grafické provedení mapy:**
 - rám mapy,
 - zeměpisná síť,
 - popis,
 - hranice států/územních celků,
 - severní Afrika, (FORECAST)
 - severní Evropa, (FORECAST)
 - Antarktida, (TEKTO)
 - stínovaný reliéf,
 - znázornění oblačnosti, (FORECAST)
 - znázornění tlaku, (FORECAST)
 - ikony.
- **Ostatní prvky:**
 - možnost zoom in/zoom out,
 - statická/dynamická.

K jednotlivým prvkům všech kategorií byla následně doplněna hodnota koeficientu, která je inspirována Likertovou stupnicí:

- 1 = odpovídá,
- 0,5 = odpovídá částečně,
- 0 = neodpovídá.

Koeficienty popisují příslušnost jednotlivých prvků z výše vypsanych kategorií v mapě a pomocí nich je možné mapy snáze popisovat a dále zpracovávat. Likertova stupnice je jednou z nejrozšířenějších stupnic v oblasti sociálních věd a je běžně rozdělována do pěti stupňů (Eskey, O'Connor 2005). V rámci této práce bylo pracováno pouze se třemi hodnotami, které byly pro hodnocení map dostačující. Pro popis jednotlivých map bylo využito znalostí kartografie získaných na Katedře geoinformatiky Přírodovědecké fakulty Univerzity Palackého.

MAPA ČÍSLO:	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Základní kompoziční prvky:														
legenda	1	0	0	1	1	1	1	0	0	0	1	0	0	0
měřítko číselné	0	0	0	1	0	0	0	0	0	0	0	0	0	0
měřítko grafické	1	0	0	1	0	0	0	0	0	0	0	0	0	0
tiráž	1	0	0	0	0	1	0	0	0	0	0	0	0	0
název: nadpis	1	1	0	1	0	1	0	0	0	0	0	0	0	0
podnadpis	0	0	0	1	0	0	0	0	0	0	0	0	0	0
mapové pole	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SOUČET_ZKP:	5	2	1	6	2	4	2	1	1	1	2	1	1	1
Nadstavbové kompoziční prvky:														
směrová růžice	0	0	0	0	0	0	0	0	0	0	0	0	0	0
zeměpisná síť	0	0	0,5	1	0	0,5	0	0	0	0,5	1	1	0	0
loga	1	0	0	0	0	0	0	0	0	0	0	0	0	0

Obr. 4.1 Ukázka tabulky části hodnocených prvků map s vyplněnými koeficienty.

Tento popis jednotlivých prvků map pomocí koeficientů (obr. 4.1) slouží k přehlednému kartografickému popisu map (příloha 1, 2 a 3). Přestože by bylo možné hodnoty koeficientů sečíst a podle bodového součtu ohodnocení v jednotlivých kritériích vytvořit shluky, které by bylo možné porovnat se shluky z embedderu Painters a shluků z uživatelského experimentu, nebyl tento krok nakonec proveden, protože nebyla kvantifikována barevnost map. Důvodem bylo nenalezení vhodného způsobu kvantifikace barevnosti. Ta se totiž na provedených experimentech provedených v rámci této práce ukázala (obr. 6.6) jako jeden z velmi výrazných prvků výtvarného stylu a její absence by výrazně ovlivnila výsledné shlukování map a následné vyhodnocení. Další možností, jak využít tabulkové podoby dat v softwaru Orange, je provedení predikce klasifikace dalších map nebo natrénování neuronové sítě. Těto problematice se blíže věnuje kapitola 9.

4.1 Dataset FORECAST

V datasetu FORECAST je obsaženo 25 online map z 11 webových stránek, které zobrazují předpověď teploty vzduchu pro území Evropy. Tyto mapy byly pořízeny snímáním obrazovky. Pokud byly součástí mapy i další prvky, které přímo nepatřily k mapě, byly ze snímku odstraněny. Mapy ze šesti zdrojů jsou mapami dynamickými, umožňující provádět posun v zobrazovaném území a transfokaci (přiblížení a oddálení). Zbýlých pět map je statických. Ze základních kompozičních prvků map chybí u všech map měřítko, a to jak číselné, tak i grafické. Kromě mapy z darkskymap.com všechny mapy obsahují legendu a název mapy. Tiráž se nachází u všech map, kromě mapy z eldoradoweather.com. V tiráži jsou nejčastěji informace, z jaké webové stránky mapa pochází, kdy byla vytvořena, nebo jaký je zdroj dat v dané mapě. V některých případech je tiráž na mapovém listu zakomponována do ikony, pomocí které se uživateli při najetí myši zobrazí informace z tiráže.

Mapy obsahují nadstavbové kompoziční prvky, především loga webových stránek nebo zdrojů dat, ze kterých mapa čerpá, nebo ikony s piktogramy o počasí. Mapa z meteoblue.com obsahuje jako jediná zeměpisnou síť. Výšeč vyznačující vedlejší mapu s podrobnějším předpovědním modelem se vyskytuje u jedné z map z yr.no.

Jednotlivé mapy se výrazně liší svým grafickým provedením. Nejvýraznější rozdíly jsou viditelné především u barevných stupnic pro teplotu vzduchu. Mapy mají také rozdílné provedení barev oceánu a moří. Ve většině případech jsou tyto plochy v barvách shodných s barvami pro teplotu vzduchu na pevnině (mapy z Windy.com, ventusky.com, maps.darksky.net, yr.no, accuweather.com, eldoradoweather.com, meteoblue.com, meteox.com com), u map z portal.chmi.cz, meteocentrale.ch a weatheronline.cz je oceán a okolní moře provedeno odlišnou barvou – odstíny modré s šedou. Další odlišnosti jsou v provedení popisu nebo výběru barvy linií. Webová stránka meteoblue.com u své mapy používá jako jediná stínovaný reliéf. Mezi další prvky, které lze v mapách nalézt patří znázornění tlaku vzduchu (meteox.com) a oblačnosti (chmi.cz).

4.2 Dataset TURIST

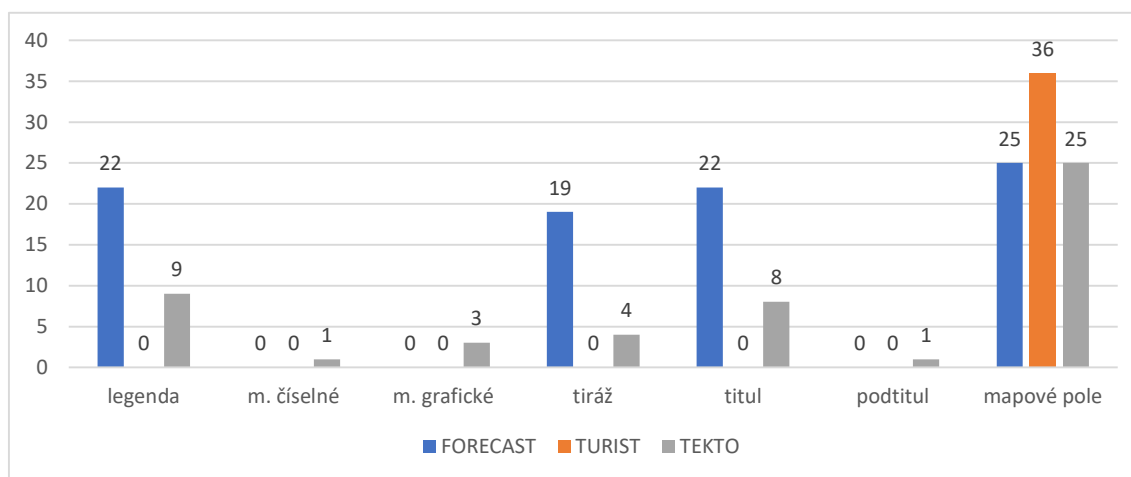
V datasetu TURIST je obsaženo 36 dynamických online map čtyř různých oblastí České republiky z devíti různých webových stránek. Tři z těchto zdrojů obsahují přímo mapy turistické, u ostatních webových stránek se jedná o podklady tematických map, jejichž obsah není striktně vymezen a díky podobnostmi s mapami turistickými vhodně doplňují dataset TURIST (viz kapitola 7).

V rámci datasetu TURIST byly pro analýzu výtvarného stylu ze snímků odstraněny základní a nadstavbové kompoziční prvky map a pracovalo se pouze s mapovým polem vybraných oblastí. Grafické provedení map v datasetu je velmi podobné a větší odlišnosti jsou viditelné mezi mapami turistickými a mapami k turistickým mapám doplněnými. První zmíněné obsahují například navíc stínovaný reliéf, ve většině případů mají také větší grafickou náplň a výraznější barvy.

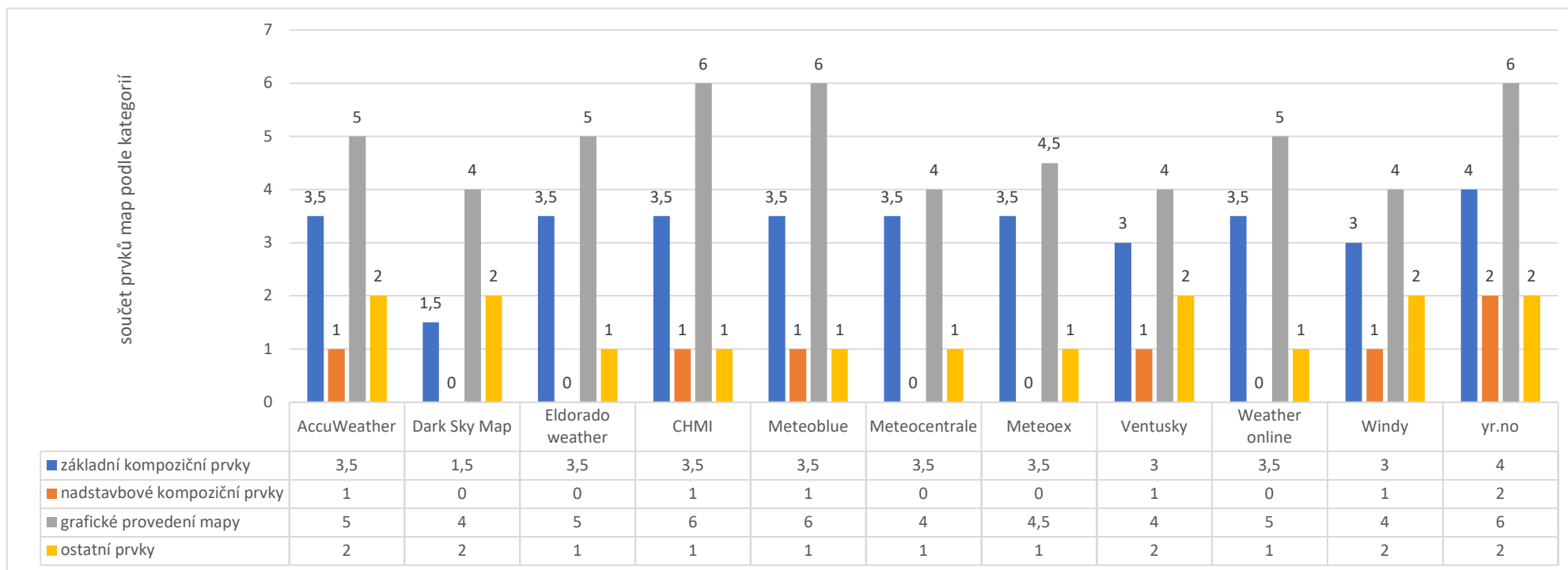
4.3 Dataset TEKTO

Třetí dataset TEKTO tvoří 25 map tektonických. Ty pocházejí z různých webových stránek. Všechny mapy jsou statické. Ze základních kompozičních prvků je pouze u třech map grafické měřítko. U šesti map nechybí v mapové kompozici název mapy. Devět map obsahuje legendu a tři mapy tiráž. Z nadstavbových kompozičních prvků se u map v datasetu vyskytují loga a zeměpisná síť (příloha 3).

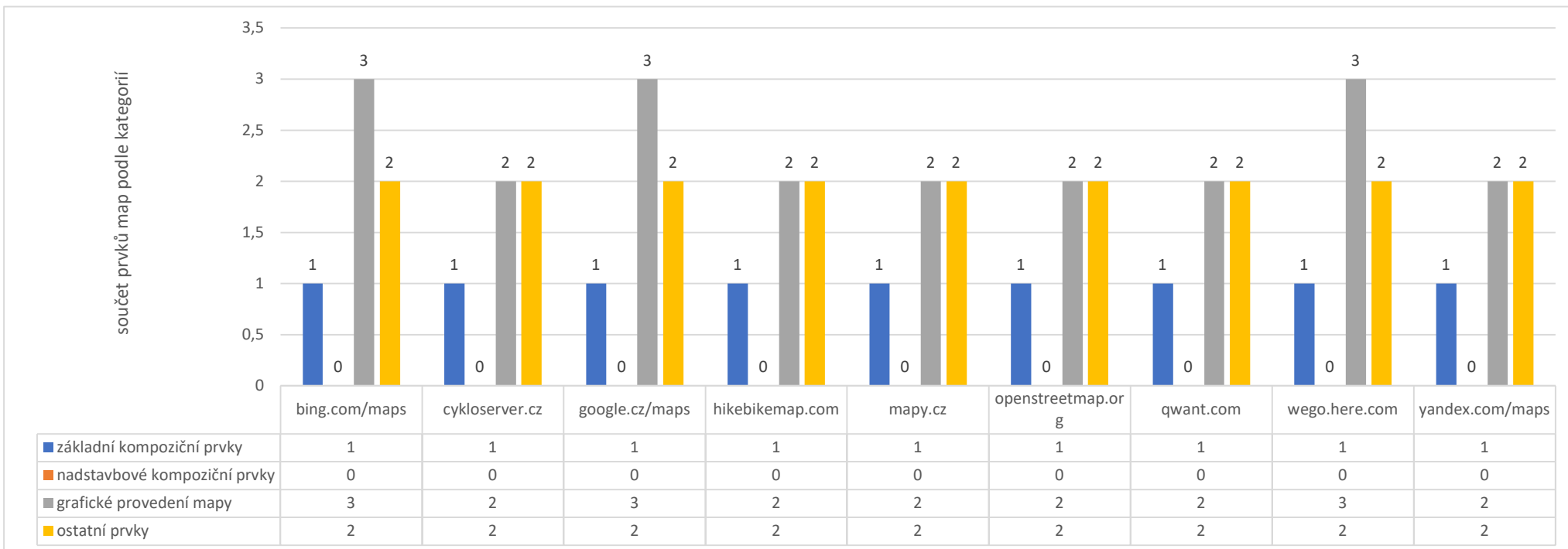
V datasetu je několik map, které neobsahují základní ani nadstavbové kompoziční prvky. V těchto mapách jsou na topografickém podkladu liniemi znázorněny pouze hranice jednotlivých tektonických desek.



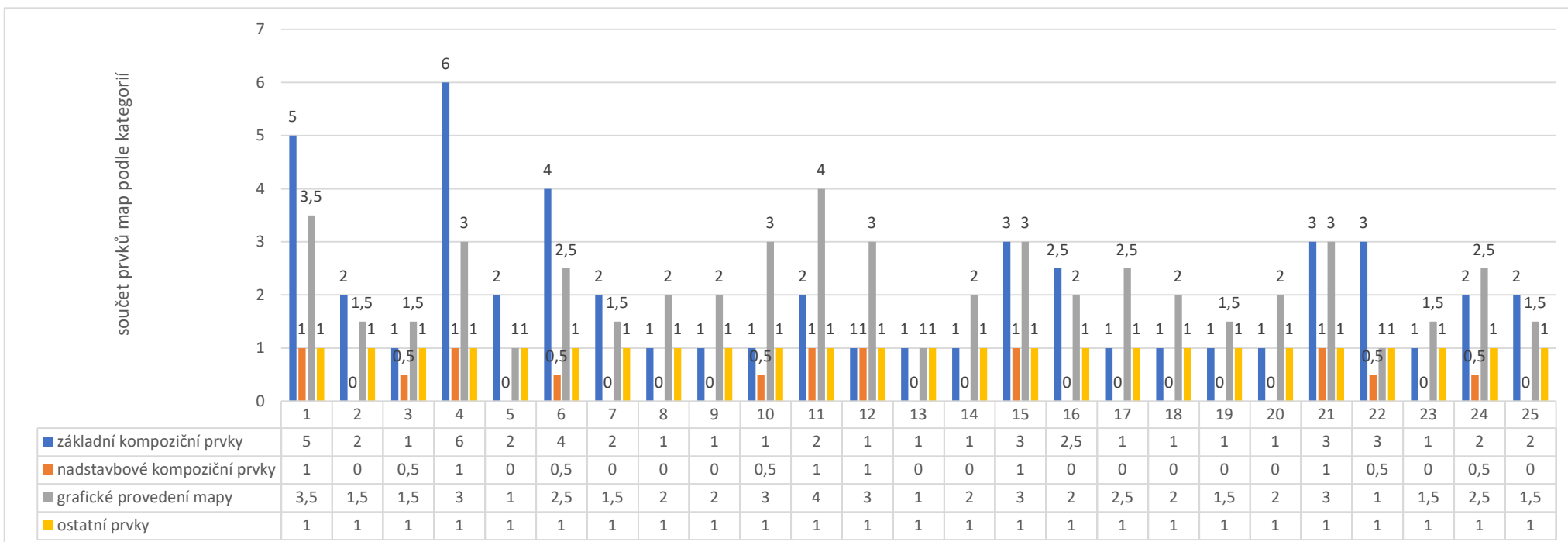
Obr. 4.2 Počet map v datasetech obsahující základní kompoziční prvky.



Obr. 4.3 Součet hodnot koeficientů jednotlivých kategorií u datasetu FORECAST podle zdroje.



Obr. 4.4 Součet hodnot koeficientů jednotlivých kategorií u datasetu TURIST podle zdroje.



Obr. 4.5 Součet hodnot koeficientů jednotlivých kategorií u map datasetu TEKTO.

5 TVORBA WORKFLOW V ORANGE

Workflow v softwaru Orange bylo sestaveno ze čtyř kroků:

Načtení snímků

Při práci s nástrojem *Image Embedder*, který umožňuje využití neuronových sítí v programu Orange, byly nejprve načteny snímky (mapy) pomocí nástroje *Import Images*. Byla zvolena složka, ze které došlo k načtení obrazových dat. Alternativou je načtení tabulky, ve které jsou odkazy na umístění obrazových dat na internetu. Nástroj *Image Viewer* posloužil k náhledu na načtené mapy.

Proces embeddování

Poté byla data napojena do nástroje *Image Embedder*. V nastavení tohoto nástroje byl vybrán atribut obrázku, který je nastavený na hodnotu *Image*, a potom samotný embedder. Na výběr je celkem sedm embedderů, jejich bližší popis je v kapitole 3.3. Čas embeddování je závislý na počtu zpracovávaných snímků. Výsledek se znázorňuje pomocí nástroje *Image Grid*, ve kterém je možné navolit matici s určitým počtem sloupců a řádků, v rámci nichž se zobrazí výsledek z nástroje *Image Embedding*. Případně je možné toto nastavení ponechat na automatickém výběru. Data je možné za účelem zobrazení v matici ořezat na čtverec nebo je zmenšit tak, aby nepřesahovala prostor buňky v matici.

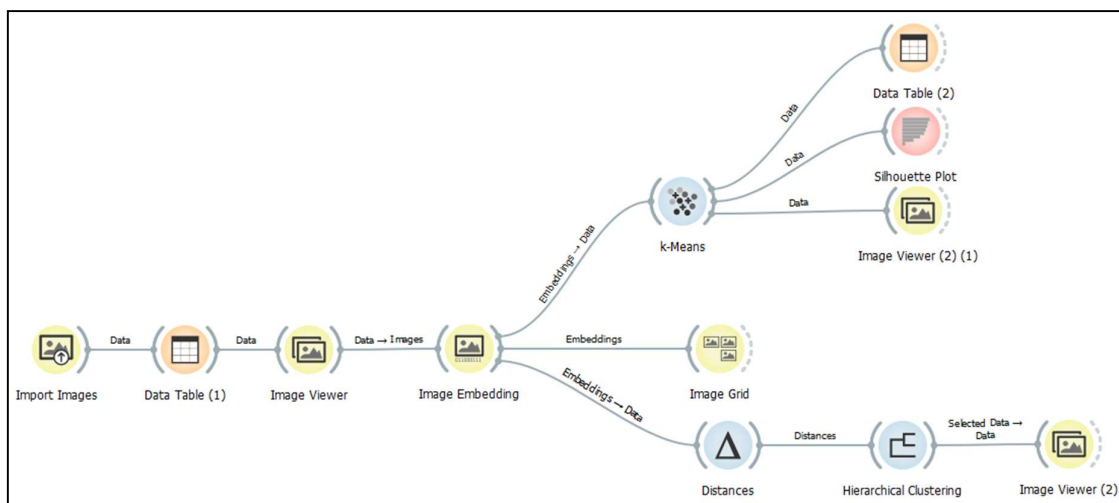
Nehierarchické shlukování

Takto vytvořená posloupnost jednotlivých nástrojů byla doplněna o nástroje z doplňku *Unsupervised*, a to o nástroje *Distances* a *Hierarchical Clustering*. Nástroj *Distances* umožňuje z polohy snímků v n-rozměrném prostoru (Dušek, Koritáková, 2017), konkrétně v prostoru, který je tvořen 2 048 rozměrů vytvořeným embedderem, vizualizovat pomocí různých vzdálenostních metrik proces hierarchického shlukování (*Hierarchical Clustering*) v podobě dendrogramu. Výčet metrik a jejich anglický popis je k dispozici na <https://orange3.readthedocs.io/> pod záložkou *Distances* (URL4). Pro následující experimenty byla využita metrika Kosinova díky jejímu úspěšnému využití v práci Dobešové (2019). Tato vzdálenost počítá vzdálenosti mezi body jako kosinus úhlu mezi dvěma vektory vnitřního produktového prostoru (URL4). V nástroji *Hierarchical Clustering* se nastavuje metrika pro měření vzdáleností mezi shluky. Na výběr je pět aglomerativních metod hierarchického shlukování, tedy metod, u kterých dochází k postupnému shlukování objektů do výsledných shluků (Řezanková a kol., 2007): metoda nejbližšího souseda (single linkage), metoda průměrné vazby (average

linkage), metoda váženého párování s aritmetickým průměrem (weighted linkage), metoda nejvzdálenějšího souseda (complete linkage) a Wardova metoda. Podle Řezankové (str. 97, 2007) Wardova metoda spojuje shluky, u kterých je přírůstek celkového vnitroskupinového součtu čtverců odchylek jednotlivých hodnot od shlukového průměru minimální. Právě tato metoda byla vybrána pro vizualizaci dat dendrogramem nástrojem *Hierarchical Clustering*. Důvodem bylo její úspěšné využití v práci Dobešové (2019).

Nehierarchické shlukování

Workflow v Orange bylo doplněno také nástrojem nehierarchického shlukování *k-Means* z doplňku *Unsupervised* a nástrojem *Silhouette Plot* z doplňku *Model* (použito v kapitole 7). Metoda *k-Means* je jednou z nejpoužívanějších shlukovacích technik, a to především díky její jednoduchosti a rychlosti. Tato metoda rozděluje data do shluků přiřazením každého objektu k těžišti jeho nejbližšímu shluku (Dean, 2014). Vstupem do nástroje *k-Means* byl stejně jako v případě hierarchického shlukování výstup z nástroje *Image Embedding*. V nástroji *k-Means* je nejprve nutné nastavit počet shluků. První možností je nastavit fixní počet shluků, druhou možností je vyhledání nejvhodnějšího počtu shluků z uživatelem definovaného rozmezí. Nástroj *Silhouette Plot* byl využit pro grafické znázornění konzistence dat ve shlucích a pro analýzu výsledků pomocí *silhouette score*. Workflow bylo také doplněno nástrojem *Data Table* pro znázornění dat v tabulce. Výsledné propojení nástrojů vytvářejících použité workflow je na obrázku 5.1.

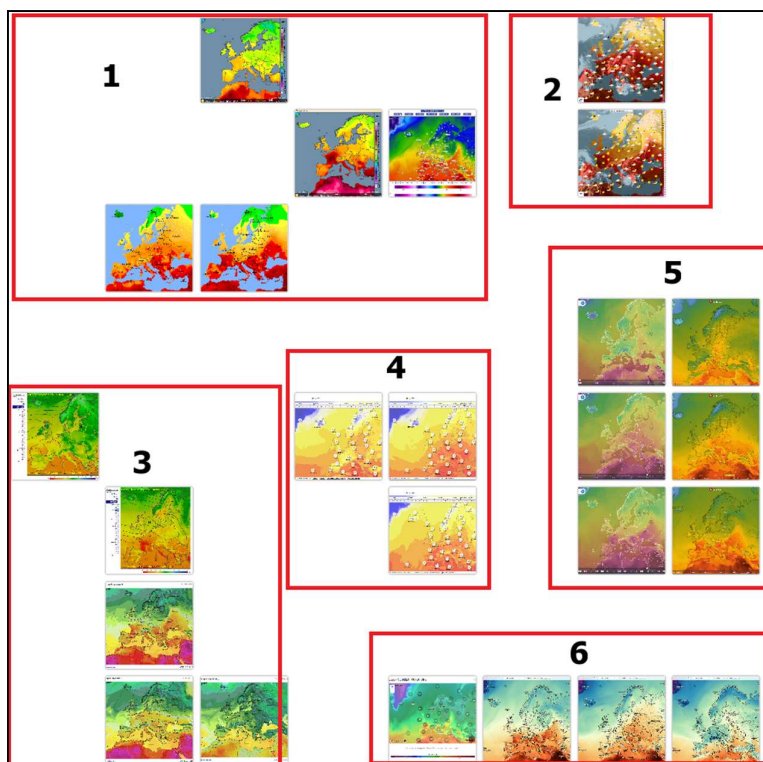


Obr. 5.1 Propojení nástrojů v Orange pro proces embeddingu a hierarchického a nehierarchického shlukování.

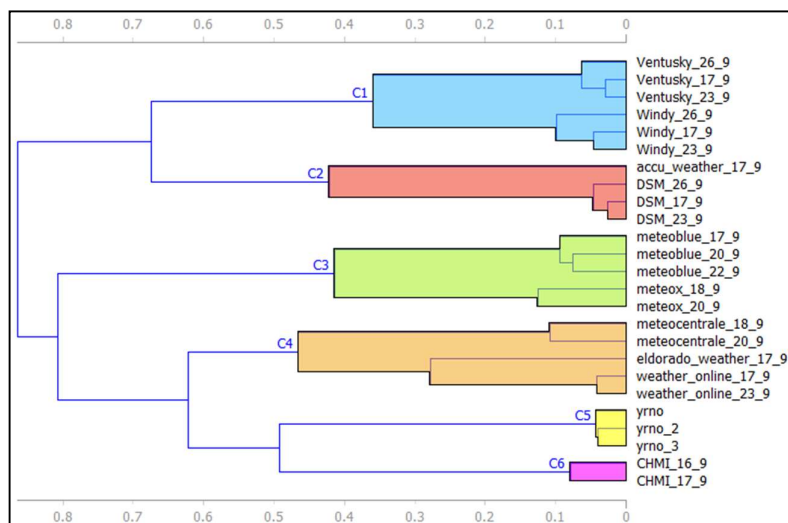
6 ANALÝZA VÝTVARNÉHO STYLU MAP PŘEDPOVĚDI POČASÍ

Pro praktické ověření fungování vytvořeného workflow v Orange bylo z webových stránek (Windy.com, ventusky.com, maps.darksy.net, portal.chmi.cz, yr.no, accuweather.com, eldoradoweather.com, meteoblue.com, meteocentrale.ch, weatheronline.cz, meteoX.com) staženo celkem 25 map předpovědi teploty vzduchu (dataset FORECAST) pro území Evropy v daném období (září 2020). Nástroj *Image Embedder* byl nastaven na neuronovou síť Painters z důvodu osvědčení tohoto embedderu při analýze kruhových výsečí měst z Urban Atlasu, který provedla ve své práci Dobešová (2019). Embedder Painters je schopen rozdělit vizuální prostor uměleckých děl na základě jejich jedinečného stylu (URL11), avšak tento algoritmus nevykazuje příliš dobré výsledky při práci s neznámými umělci. Natrénován byl na umělecká díla, hodnotí a analyzuje spíše umělecký styl, nikoliv výtvarný styl. Proto byla použitelnost embedderu Painters otestována vizuálním porovnáním výstupů ze všech poskytovaných embedderů softwaru Orange (obr. 3.3). Právě u výsledků vytvořených embedderem Painters znázorněných pomocí matice je nejviditelnější „shlukování“ map ze stejného zdroje. Výstupem procesu při znázornění v matici nástrojem *Image Grid* je **šest** viditelných skupin (obr. 6.1). Shluky jsou viditelné také při vizualizaci procesu shlukování pomocí dendrogramu (obr. 6.2). Při znázornění prvních šesti shluků tyto shluky přesně odpovídají shlukům viditelným pouhým pozorováním výsledků znázorněných maticí (obr. 6.1). Při porovnání výsledků pomocí dendrogramu s nastavením 11 shluků, tedy pro každý zdroj jeden shluk, dosahovaly stejných výsledků jako embedder Painters i ostatní embeddery. V každém shluku byly pouze mapy z jednoho zdroje. Jedinou výjimkou byl embedder Inception v3, u kterého byl problém se zařazením jedné mapy z meteoblue.com a mapy z eldoradoweather.com.

Při rozboru výsledné matice 25 map teploty vzduchu Evropy z embedderu Painters je patrné, že došlo ke shlukování map ze stejných webových stránek a ke spojení těchto dvojic, případně trojic do shluků, a to především na základě barevnosti map. Tímto výsledkem byl potvrzen potenciál tohoto embedderu pro vymezení výtvarných stylů map. Shluky jsou tvořeny mapami z jednotlivých webových stránek a mají tedy jednotný výtvarný styl. Viditelný je také výrazný vliv barev, který byl demonstrován experimentem s barevnými mapami a mapami v odstínech šedi popsány níže (kapitola 6.1).



Obr. 6.1 Šest shluků vytvořených neuronovou sítí Painters znázorněných pomocí nástroje *Image Grid*.

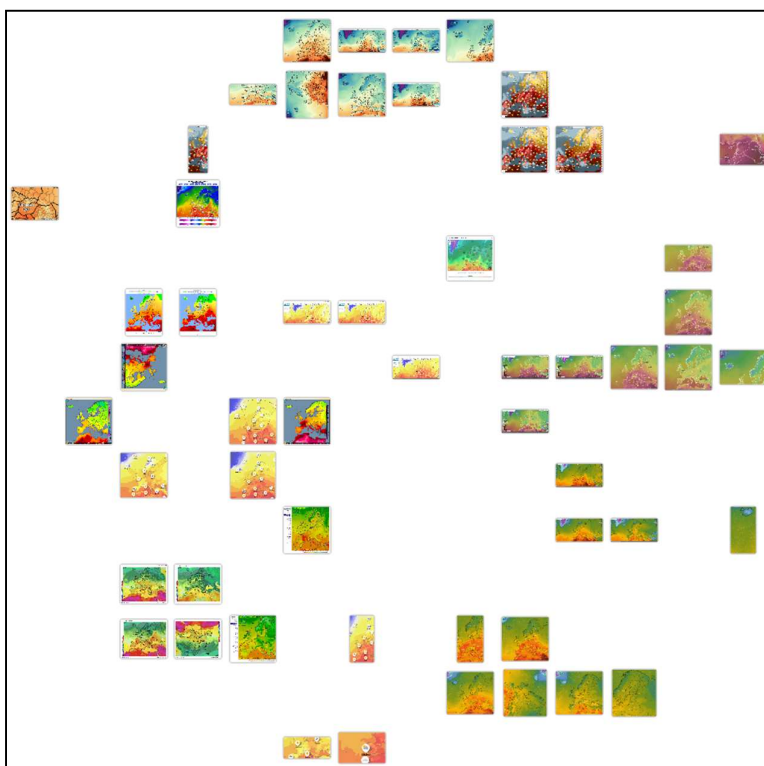


Obr. 6.2 Znázornění shluků originálních dat datasetu FORECAST z hierarchického shlukování získaných pomocí embedderu Painters.

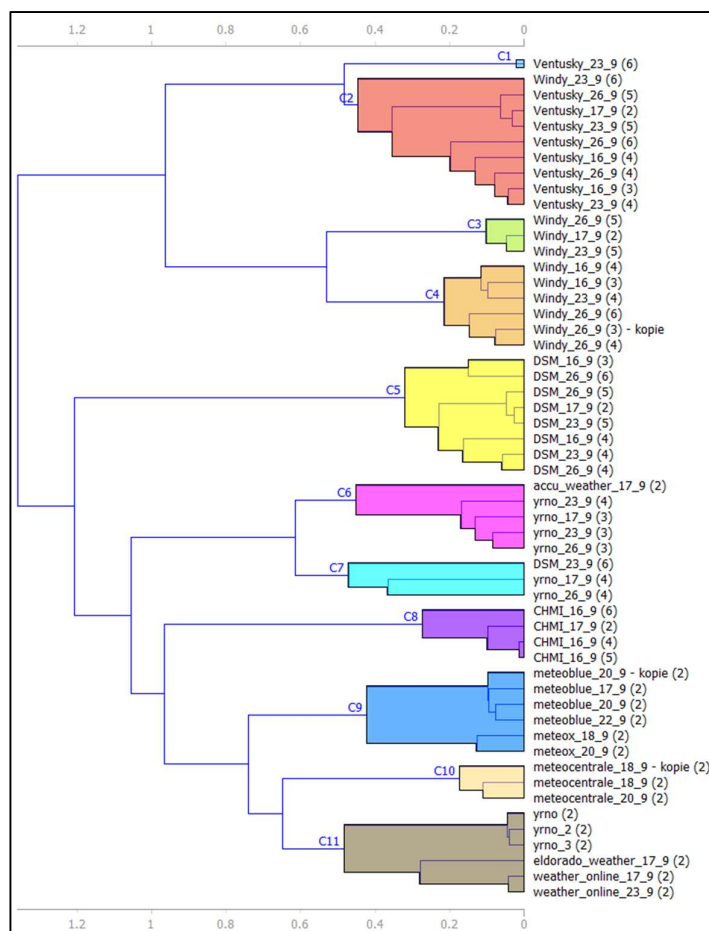
Na výše uvedených analyzovaných mapách a na jejich derivátech (ořez, otočení, změna rozlišení) byly následně provedeny experimenty popsané v podkapitolách zkoumající **barevnost** (podkapitola 6.2) jako jeden z hlavních parametrů výtvarného stylu a parametry vstupních dat s potenciálem ovlivnit proces shlukování – **ořez**, **otočení** (podkapitola 6.1) a **rozlišení** (podkapitola 6.3).

6.1 Ořez a otočení vybraných map

Nejprve bylo zjišťováno, jaký efekt na vytvoření shluků má přidání některých **výřezů** již použitých map a **otočení** náhodně vybraných map do předchozího mapového datasetu FORECAST. Navýšením počtu vstupních dat (z 25 na 54) se shluky embedderem Painters staly méně kompaktní. Proces jejich vzniku je totožný s předchozím krokem. Pravděpodobně v důsledku použití kombinace většího množství dat, výřezů a pootočení map není zařazení některých výřezů map k originálním mapám (především s méně výraznými barvami) úspěšné (obr. 6.3). Při zobrazení výsledků pomocí dendrogramu s 11 shluky (obr. 6.4) je viditelný problém u výrazných výřezů z map Ventusky (mapa s názvem Ventusky_23_9 (6)), Windy (mapa s názvem Windy 23_9 (6)), DarkSkyMap (mapa s názvem DSM_23_9 (6)) a yr.no (mapy s názvem yr_no_17_9 (4) a yr_no_26_9 (4)). V těchto případech došlo k výřezu asi 6 % celé mapy. Pokud se jednalo o větší část originální mapy (například pouze čtvrtina originální mapy) byl tento výřez přiřazen správně do shluku map ze stejného zdroje.



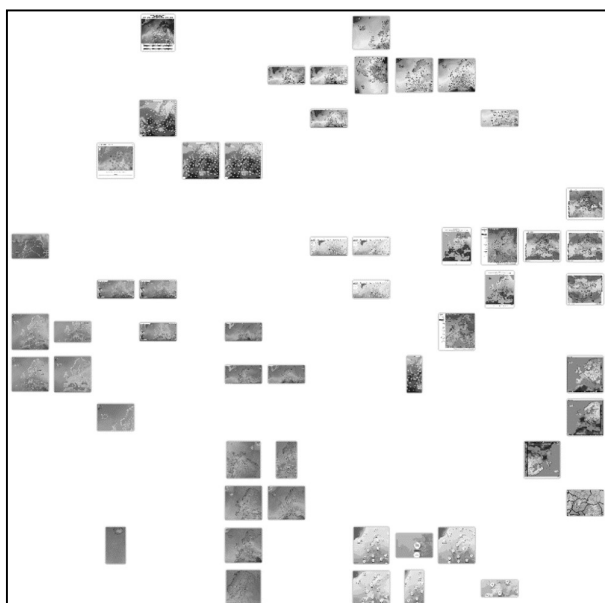
Obr. 6.3 Kombinace výchozích map s jejich výřezy a otočenými vybranými mapami.



Obr. 6.4 Dendrogram 11 shluků originálních map s vybranými ořezanými a otočenými mapami z originálního datasetu.

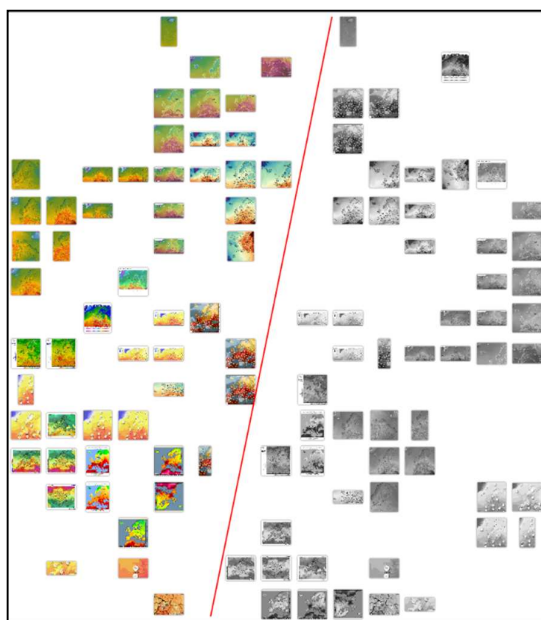
6.2 Barevnost map

Barevnost map je podle definice jedním z parametrů výtvarného stylu a vzhledem k výraznému ovlivnění zařazení map do shluků právě na základě barevného provedení map byly všechny mapy z předchozí části (podkapitola 6.1) převedeny do **odstínů šedi** v softwaru Zoner Photo Studio 14 funkcí *Odstíny šedi*. Proces embeddingu je výpočetně náročnější v porovnání s původními, barevnými mapami, i přes stejnou velikost vstupních dat (celkem 54 map). U originálních map trval 4,5 sekundy, u map v odstínech šedi 5,25 sekundy. Výsledné shluky jsou při porovnání se shluky barevných map odlišné (obr. 6.3 a 6.5). Důvodem může být xeroxový efekt, tedy jev, ke kterému dochází při převodu map do černobílé podoby, důsledkem čehož mohou dva odlišné barevné tóny vytvořit stejný odstín šedé barvy (Voženílek, Kaňok, 2011).



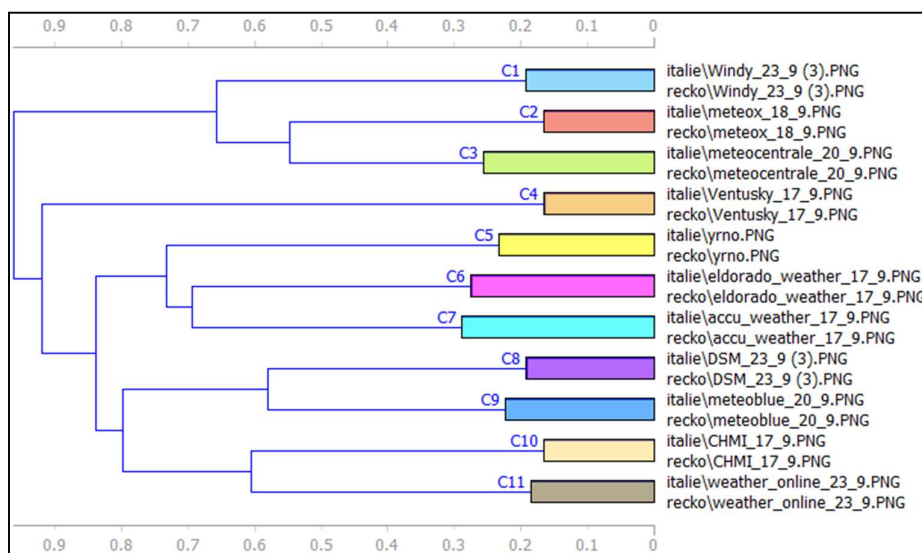
Obr. 6.5 Převedení map z kapitoly 6.1 do odstínů šedi.

Následně byla provedena **kombinace** dvou výše zmíněných kategorií, tedy map s originální barevností a jejich výřezy s otočenými variantami, a mapy v odstínech šedi (celkem 108 map nebo jejich částí). Tato část experimentu měla za cíl zjistit, zda bude docházet k seskupování map ze stejných zdrojů i přes rozdílnou barevnost map. Mapy v odstínech šedi jsou totiž seskupené v pravé části gridu, mapy barevné jsou naopak pouze v levé části (obr. 6.6). Nedošlo tedy k žádnému promíchání map mezi sebou. Z tohoto vyplývá velký vliv barevnosti na výsledné zařazení map.



Obr. 6.6 Kombinace barevných map a map v odstínech šedi.

Vliv barevnosti byl potvrzen také při vizualizaci výřezů barevných variant map dvou oblastí (obr. 6.7), a to Řecka a Itálie pomocí dendrogramu (celkem 22 map). Výsledkem je vznik dvojic shluků ze stejných map. Viditelné je tedy upřednostnění barevnosti map před strukturou mapového pole, tedy tento fakt potvrzuje předpoklad možného využití tohoto embedderu pro definování výtvarných stylů map. Pokud by došlo k upřednostnění struktury mapového pole, vznikaly by spíše shluky, ve kterých by byly obsaženy shodné oblasti.

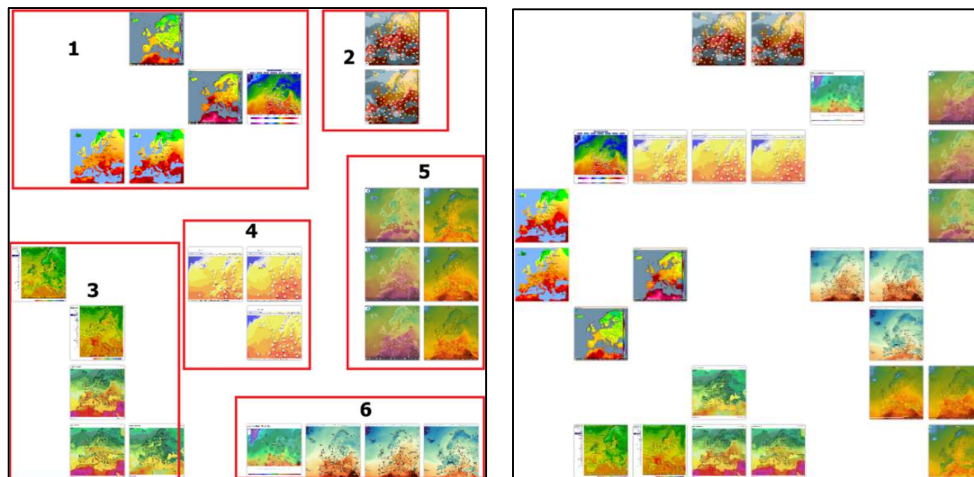


Obr. 6.7. Dendrogram dvojic map dvou oblastí.

6.3 Rozlišení map

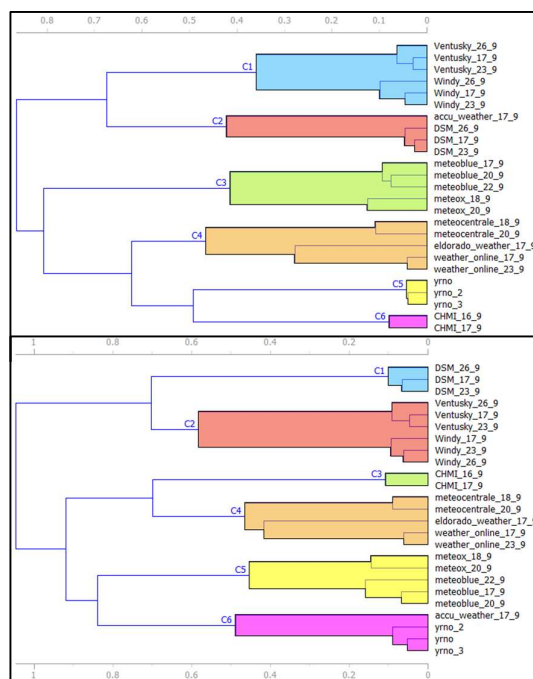
Pro správné naučení a fungování embedderu Painters se vstupní data upravují na rozměry 256 x 256 pixelů. Každý snímek vstupující do neuronové sítě pro její trénování byl autorem sítě zmenšen na kratší straně na 256 pixelů a poté ořezán ve středu delší strany tak, aby bylo dosaženo požadovaného rozměru. V případě zpracování jakéhokoliv datasetu s odlišným datovým **rozlišením** vstupních dat by mohlo mít rozdílné rozlišení vliv na výsledné zařazení jednotlivých map do shluků. Mapy základního datasetu byly proto zmenšeny podle instrukcí výše a byl na ně aplikován embedder Painters. Problematické je ořezání vstupních dat na čtvercový formát a s tím související ztráta informace v mapách. Čím je poměr stran odlišnější od čtvercového formátu, tedy poměru stran 1:1, tím je ztráta informace větší. Ke ztrátě informace dochází také v důsledku zmenšení vstupních map na požadované rozměry. Díky tomu je nadhodnoceno barevné provedení map, zatímco kompozice nebo barevné provedení hranic či dalších méně výrazných prvků na mapách je potlačeno. Při vizuálním porovnání (obr. 6.8) výsledných shluků se shluky

originálních map je viditelný rozdíl ve výsledcích. Není zde jednoduché provést rozdělení matice na šest shluků, jako v případě tvorby shluků s originálními mapami.



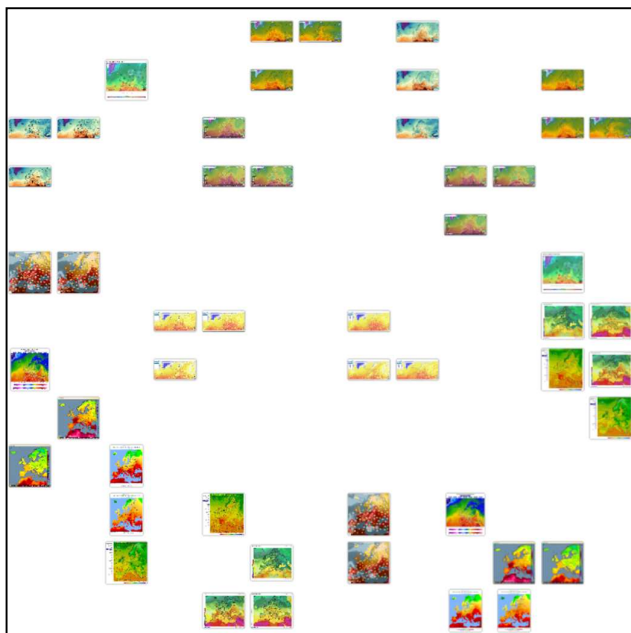
Obr. 6.8 Rozdíly ve výsledcích shlukování při použití originálních dat bez provedení změny rozlišení a poměru rozměrů (vlevo) a s normalizací těchto parametrů (vpravo).

Proto bylo využito možnosti procesu shlukování pomocí dendrogramu (spodní část obr. 6.9). Po nastavení zvýraznění šesti shluků je jediným rozdílem přidělení mapy z portálu accuweather.com mezi mapy z yr.no namísto mapy z maps.darksky.net, jak tomu bylo v případě shlukování u originálních map. Dalším rozdílem, viditelným na dendrogramu, je odlišný proces tvorby těchto shluků (obr. 6.9).



Obr. 6.9 Odlišnosti při vizualizaci procesu shlukování pomocí dendrogramu s originálními daty (nahore) a s daty zmenšenými a oříznutými na rozměr 256 x 256 px.

Další částí experimentování s rozlišením bylo zmenšení originálních map na 10 % své velikosti a spolu s originály byly znovu importovány do softwaru Orange, kde na ně byl použit embedder Painters. Stejně mapy s odlišným rozlišením vytvořily vlastní shluky, které v mnoha případech ani nebyly v blízké vzdálenosti (obr. 6.10). Z výše uvedeného je viditelné, že vliv na výsledek má tedy nejen barevnost, ale i rozlišení vstupních map. Oproti barevnosti, která je jednou z charakteristik výtvarného stylu, je vliv rozlišení na výsledek problematický. Pro hlavní část práce je možné mapy transformovat do poměru stran 1:1 a změnit jejich rozlišení, tak, aby všechny mapy byly stejné. Krok normalizace vstupních dat má nicméně za následek ztrátu informace u některých map a snížení vlivu méně výrazných částí map a tím zvýšení vlivu barevného provedení. Druhou možností je mapy nenormalizovat a ponechat jejich rozlišení a poměry stran jako jeden z parametrů definujících výtvarný styl map.



Obr. 6.10 Výsledky experimentu se zmenšením vybraných map na 10 % originální velikosti (viditelné jsou shluky stejných map s normálním a se sníženým rozlišením).

6.4 Výsledky experimentů

Z výsledků provedených experimentů s mapami teploty vzduchu (dataset FORECAST) s využitím embedderu Painters vyplývají následující poznatky:

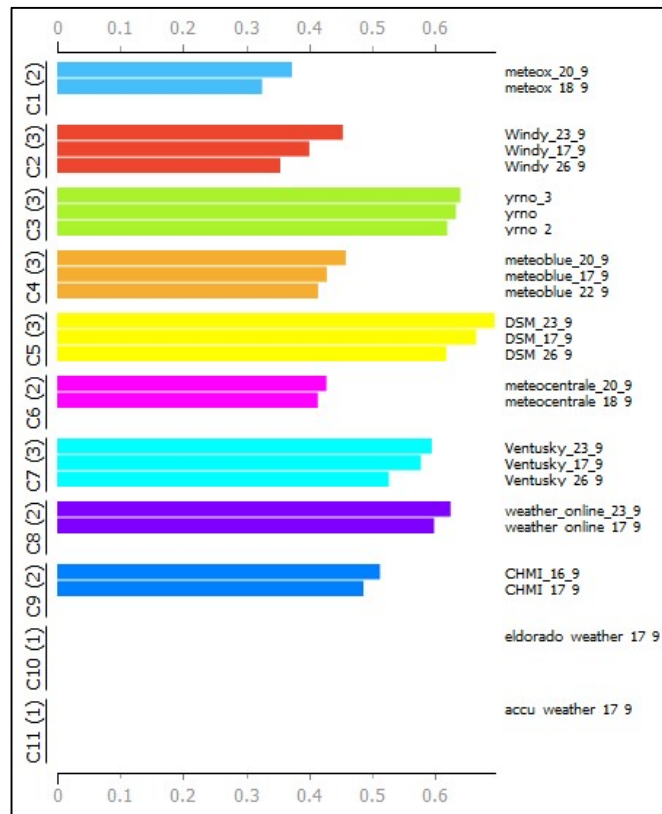
1. Nejvýraznější vliv na shlukování map byl prokázán u jejich **barevnosti** (obr. 6.6, obr. 6.7):
 - Vliv barevnosti na shlukování je prokazatelný již při shlukování originálních map (obr. 6.1 a 6.2) i při experimentu s dílčími oblastmi z originálních map (obr. 6.7).
 - Velmi výrazný vliv barevnosti byl prokázán při experimentu s mapami převedenými do odstínu šedi (při rozdělení vstupního datasetu na dvě poloviny) (obr. 6.6).
 - Vliv barevnosti mapy na shlukování potvrzuje potenciál využití neuronových sítí, konkrétně embedderu Painters pro definování výtvarných stylů.
2. Vliv na shlukování byl prokázán u **rozlišení** map
3. a **výřezů** z originálních map (kapitoly 6.1 a 6.3):
 - Tvorba výřezů je problematická pouze při použití značných výřezů (méně než 6 % mapy) z originálních map (obr. 6.3). Pro finální shlukování map však nebyl důvod provádět takto velké výřezy (využity byly buď celé mapy nebo došlo k větším výřezům) a není tedy problematiku výřezu nutné více zvažovat.
 - Různé rozlišení snímků (map) v důsledku rozdílné velikosti originálních map sice může způsobit odlišné zařazení dané mapy do shluku (obr. 6.8 a 6.9), pokud ale není mapa zmenšena úmyslně (obr. 6.10), neovlivňuje tento fakt vymezení výsledných shluků.
 - Pro správné fungování embedderu Painters se doporučuje normalizovat poměr stran a rozlišení map, jak je uvedeno v kapitole 6.3. Avšak tím se ztratí některé informace v mapě, protože dojde ke zmenšení map a zároveň k jejich ořezu. Méně výrazné prvky mapy mohou být výrazně potlačeny.
4. Změna **orientace map** nemá vliv na výsledné zařazení map do shluků:
 - Při experimentu se změnou orientace vybraných map nebylo pozorováno výraznějších problémů se zařazením těchto map k jejich neotočeným variantám (obr. 6.3).

6.5 Využití nehierarchického shlukování

Jako zástupce nehierarchické metody shlukování byl využit nástroj *k-Means*, využívající stejnojmennou metodu, u které je nejprve nutné nastavit počet shluků. Při zpracování datasetu FORECAST byla zvolena možnost vyhledání nejvhodnějšího počtu shluků z uživatelem definovaného rozmezí, které bylo nastaveno od dvou do 11 shluků. Na základě získaného skóre (*Silhouette Scores* = 0,395) bylo vyhodnoceno, že nejvhodnějším počtem shluků je pro tento dataset 11. Tento počet odpovídá počtu zdrojů map a jedná se o další důkaz potvrzující potenciál embedderu pro vymezení výtvarného stylu map. Dalším krokem před spuštěním nástroje bylo nastavení inicializační metody, tedy způsobu, jakým algoritmus začíná shlukování. Na výběr jsou metody *k-Means ++* (první střed je vybrán náhodně, následující jsou vybrány ze zbývajících bodů s pravděpodobností úměrnou druhé mocnině vzdálenosti od nejbližšího středu) anebo náhodná inicializace (shluky jsou nejprve přiřazeny náhodně a poté jsou aktualizovány dalšími iteracemi) (URL20). V tomto případě byla použita metoda *k-Means ++* a Kosinova metrika. Výsledek shlukování je v tomto případě možné vyhodnotit i číselně, a to pomocí nástroje *Silhouette Score*. Tato metrika slouží k vyhodnocení kvality techniky shlukování. Jeho hodnota se pohybuje od 1 do -1 (Bhardway, 2020):

- 1: shluky jsou od sebe dobře oddělené a jasně rozlišené, prvek je velmi kompaktní v rámci shluku, do kterého patří, a zároveň daleko od ostatních shluků,
- 0: shluky se překrývají a vzdálenost mezi shluky není významná, prvky jsou na hranici mezi dvěma shluky,
- -1: shluky jsou vytvořeny nesprávně.

Výsledek z nástroje *Silhouette Score* je na obrázku 6.11. Průměrné skóre dosáhlo hodnoty **0,47**, maximální hodnota dosahuje **0,64**, což dokládá, že shluky jsou poměrně dobře oddělené. Ve dvou případech (eldoradoweather.com a accuweather.com) bylo dosaženo skóre **nulové** hodnoty. U těchto shluků dochází k překryvům a jsou na hranici mezi shluky. Z výsledků nehierarchického shlukování je dobře viditelné shlukování podle zdrojů jednotlivých map, tedy podle výtvarného stylu.



Obr. 6.11 Výsledky nehierarchického shlukování datasetu FORECAST z nástroje *Silhouette Score*.

7 ANALÝZA VÝTVARNÉHO STYLU TURISTICKÝCH MAP

V České republice se zabývá tvorbou tištěných turistických map šest vydavatelství (Gabryš, 2019):

- Geodézie On Line,
- Kartografie PRAHA,
- Klub Českých Turistů,
- Machovský,
- SHOCart,
- Žaket.

Turistické mapy lze nalézt také v online podobě, a to na stránkách Mapy.cz, OpenStreetMap Hike & Bike Map, Cykloserveru a PhoneMaps. Jak je uvedeno v kapitole 2.3 *Data*, pro získání map byly upřednostněny online zdroje. Důvodem byla snaha minimalizovat negativní vlivy při skenování map v papírové podobě jako je například vznik stínů v místech přehybů a podobně. Vstupní data byla tedy získávána vytvořením screenu požadované oblasti. I takový sběr dat má ale svá omezení, jedná se hlavně o nemožnost provedení výřezu mapy ve stejném měřítku z důvodu odlišných zoom levelů různých mapových serverů a také o nemožnost výběru vždy úplně stejného území.

Bohužel neexistuje pro Českou republiku velké množství online dostupných turistických map (pouze tři), v úvahu tedy přicházejí mapy „**topografické**“ (podklady tematických map/tematické mapy pro orientaci v terénu či pro řidiče; obsah těchto map není striktně vymezen). Vybraná území nemohou být vzhledem k limitované velikosti monitoru v příliš velkém měřítku a tak na turistických mapách nejsou turistické trasy příliš výrazné a využití topografických map se tedy velmi nabízí. Těch je dostupných online větší množství a v kombinaci se čtyřmi vybranými územími je výsledkem uspokojivý počet map vstupujících do následné analýzy.

Získání screenu požadované oblasti bylo obtížné především z důvodu překrytí mapového pole jinými prvky mapy nebo dalšími prvky, které se vyskytují v internetovém prostředí. Jedním z takových prvků byla i reklama v případě webu mapquest.com. Z toho důvodu byly screeny z tohoto webu jiné, než screeny z ostatních webových stránek. Problém nastal také u webu viamichelin.com. Nastavené zoom levely neumožňují znázornění Olomouce a okolí tak jako u ostatních stránek a proto bylo možné pracovat pouze s menším územím, kde podrobnost

vykreslených prvků nejvíce odpovídala ostatním mapám. Pro zachování nehomogennějšího území ve stejném měřítku byly tyto dva zdroje vynechány. Celkem bylo využito těchto devět zdrojů a to:

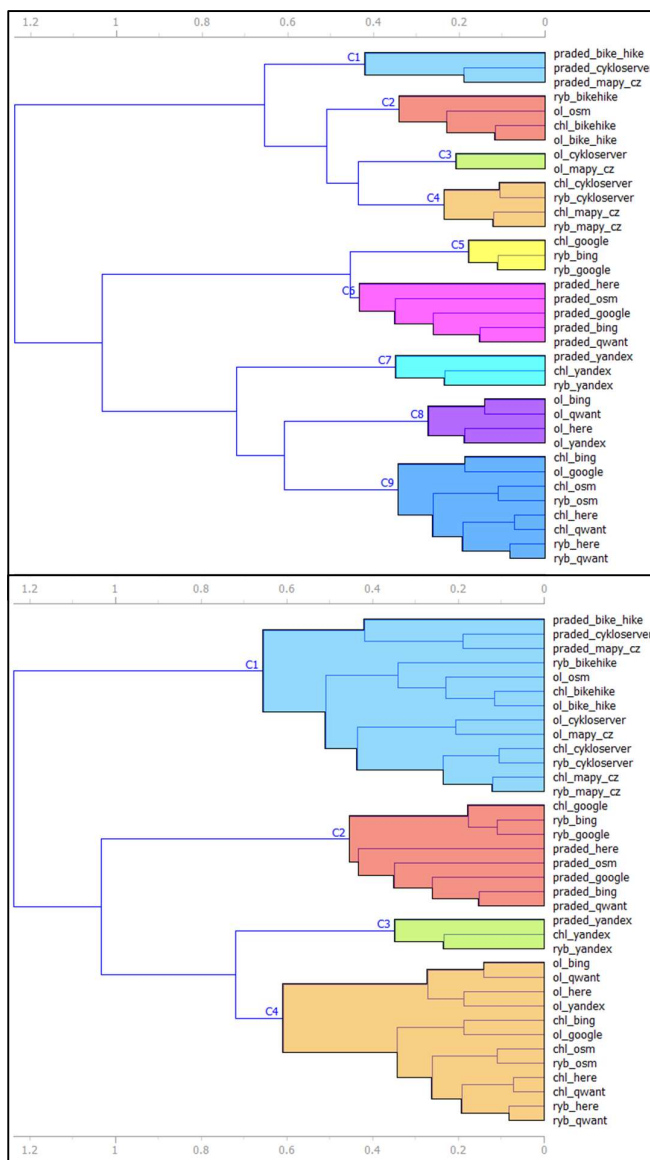
- bing.com/maps,
- cykloserver.cz,
- google.cz/maps,
- hikebikemap.com,
- mapy.cz,
- openstreetmap.org,
- qwant.com,
- wego.here.com,
- yandex.com/maps.

Z každého zdroje byl pro všechna vybraná území (viz kapitola 2.5 Zájmová území) vytvořen screen oblasti. Celkem bylo získáno 36 různých screenů. Před spuštěním procesu embeddingu byly uvažovány tři možné varianty výsledného shlukování, a to:

- dojde ke shlukování podle území,
 - embedder Painters není vhodný pro definování výtvarného stylu map
- dojde ke shlukování podle zdrojů map,
 - embedder Painters je vhodný pro definování výtvarného stylu map
 - každý zdroj má svůj vlastní výtvarný styl; případně má více zdrojů shodný výtvarný styl a tyto zdroje budou součástí jednoho shluku
- shlukování bude ovlivněno jak zdrojem, tak územím a výsledné shluky nebudou jasně diferenciovány.
 - embedder může být použitelný pro definování map na základě barevnosti, struktury, náplně mapy nebo dalších parametrů, nikoliv na základě výtvarného stylu

Výsledné shluky spadají do třetí skupiny. Nedošlo totiž k jasné diferenciaci podle zdroje nebo podle území a při znázornění výsledků dendrogramem (čtyři a devět shluků) je patrné promíchání dat z různých zdrojů a také různých oblastí v jednom shluku (obr. 7.1).

Výjimku tvoří při vizualizaci pomocí devíti shluků mapy z yandex.com a hikebikemap.org. V těchto případech tři ze čtyř map ze stejného zdroje vytvořily jeden shluk. V případě hikebikemap.org je ve shluku ale také jedna mapa z openstreetmap.org. V případě yandex.com je chybějící mapou mimo tento shluk mapa Olomouce, u hikebikemap.org se jedná o mapu s okolím Pradědu.



Obr. 7.1 Výsledky shlukování turistických map při zobrazení čtyř (spodní část) a devíti (horní část) shluků.

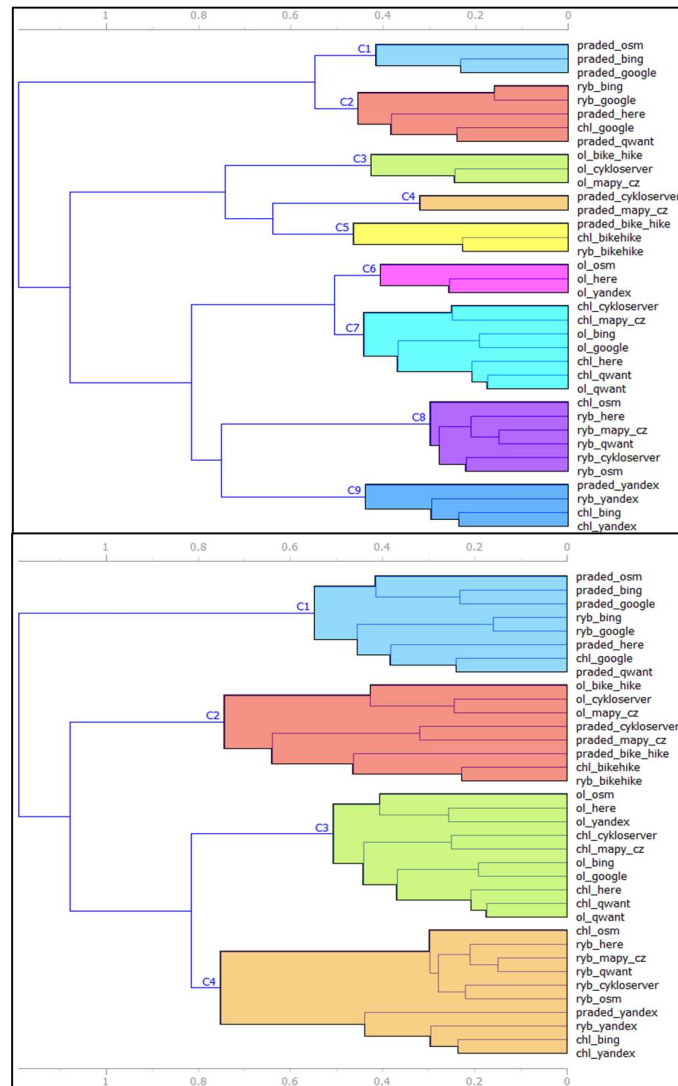
Protože by měla být pro správné fungování embedderu Painters data transformována do rozměrů 256 x 256 pixelů, došlo v dalším kroku k výřezu části v právě takovýchto rozměrech z původních dat. Z originálních dat byla vybrána tato území:

- Praděd,
- rybníky Velký a Malý Tisý,

- historické centrum Olomouce a blízké okolí,
- obec Kosičky a Třesický rybník.

Následně došlo k jejich zpracování stejně jako v předchozím kroku. Nicméně ani v tomto případě nedošlo k jasné diferenciaci podle zdroje nebo podle území a je patrné promíchání dat z různých zdrojů a také různých oblastí v jednom shluku (obr. 7.2).

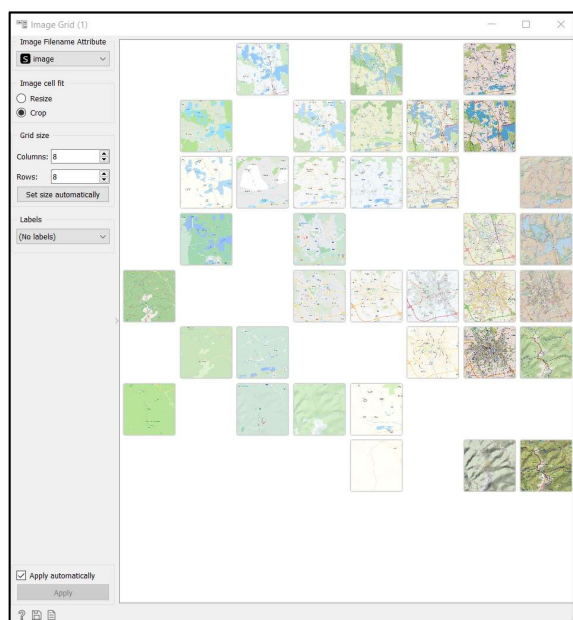
Nejúspěšnější bylo znovu shlukování výřezů z yandex.com a hikebikemap.org při znázornění pomocí devíti shluků. V těchto případech jsou tři ze čtyř výřezů v jednom shluku. I zde je součástí jednoho ze shluků mapa z jiného zdroje a to mapa z bing.com ve shluku map z yandex.com.



Obr. 7.2 Výsledky shlukování výřezů turistických map v rozlišení 256 x 256 pixelů při zobrazení čtyř (spodní část) a devíti (horní část) shluků.

7.1 Nehierarchické shlukování datasetu TURIST

Na stejném datasetu turistických map (dataset bez změny rozměrů a před zmenšením) bylo následně provedeno shlukování pomocí nehierarchických shlukovacích metod. Cílem bylo zjistit, zda nehierarchická metoda bude poskytovat lepší nebo naopak horší výsledky. Konkrétně byl využit nástroj *k-Means*, využívající stejnojmennou metodu, u které je nejprve nutné nastavit počet shluků. Při zpracování datasetu turistických map bylo využito možnosti vyhledání nejvhodnějšího počtu shluků z uživatelem definovaného rozmezí, které bylo nastaveno od dvou do devíti shluků. Na základě získaného skóre (*Silhouette Scores* = 0,107) bylo vyhodnoceno, že nejvhodnějším počtem shluků jsou pro tento dataset čtyři. Proces shlukování byl otestován s oběma inicializačními metodami.

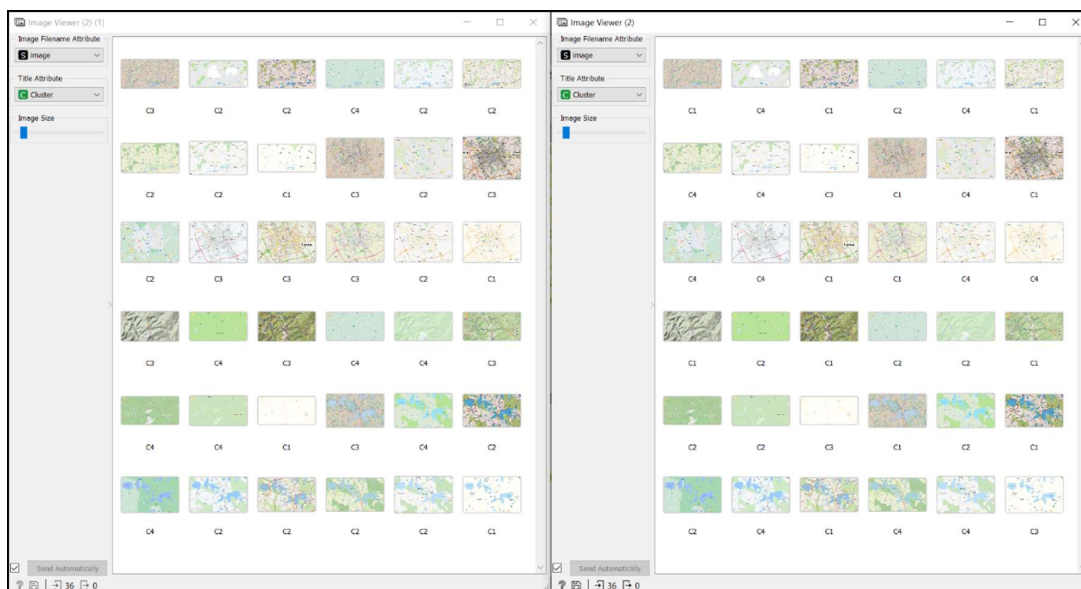


Obr. 7.3 Výsledná matice turistických map z nástroje *Image Embedding*.

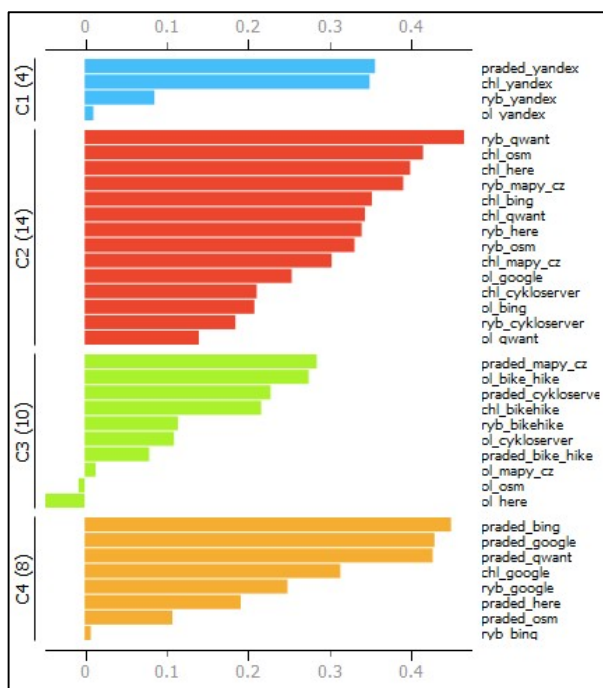
Rozdíly jsou patrné při porovnání příslušnosti jednotlivých map do shluků (obr. 7.4). Viditelné je výrazné promíchání jednotlivých map ve shlucích. Ani v jednom případě nedošlo k jasné diferenciaci podle území nebo podle zdroje. Z tohoto lze znovu usuzovat, že embedder Painters a některá z metod shlukování nejsou schopny provést shlukování map na základě jejich výtvarného stylu. Z matice na obrázku 7.3 je viditelné shlukování na základě barev, struktury a také grafické náplně map. Právě náplň map a její vliv na shlukování je zkoumána v následující kapitole 8.

Protože se matice z nástroje *Image Grid* generuje na základě výstupu z nástroje *Image Embedding*, nedochází ke změně pozic jednotlivých map v matici při použití odlišných metod shlukování. Hodnocení na základě matice tedy není příliš přesné a určení shluků jen vizuálním zhodnocením matice je velmi subjektivní a nemusí

odpovídat reálným shlukům. Pro určení shluků v případě nehierarchického shlukování je lepší využít znázornění příslušnosti map do shluku nebo v případě hierarchického shlukování provést znázornění pomocí dendrogramu.



Obr. 7.4 Rozdílné zařazení map do jednotlivých shluků. Vlevo výstup z nehierarchické metody k-Means, vpravo výstup z hierarchické metody (kosinova vzdálenost, Wardova metoda). C s příslušným číslem pod mapou označuje výsledný shluk.



Obr. 7.5 Výstup z nástroje *Silhouette Score* porovnávají kvalitu zařazení map do shluku datasetu TURIST (inicializace k-Mean ++, Kosinova vzdálenost).

Výsledek z nehierarchického shlukování datasetu TURIST byl vyhodnocen také pomocí *silhouette score*, které je výstupem stejnojmenného nástroje (obr. 7.5). Pracováno bylo s k-Means ++ inicializací a také s náhodnou inicializací. Také byly porovnány hodnoty pro všechny vzdálenosti, které je možné nastavit v nástroji *Silhouette Score*, a to Euklidovskou, Manhattanskou a Kosinovu vzdálenost (tab. 7.1). Průměrné hodnoty *silhouette score* se pohybovaly v rozmezí od 0,1 po 0,24. Není tedy dosaženo jasného vymezení shluků a některé mapy, s hodnotou blížící se nule, jsou na hranici mezi dvěma shluky. Nejlepších výsledků bylo dosaženo při použití inicializace k-Means ++ s Kosinovou vzdáleností.

Tab. 7.1 Průměrné hodnoty *silhouette score* pro čtyři shluky datasetu 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

7.2 Výsledky experimentů

Na základě dosažených výsledků bylo potvrzeno, že nejvhodnější vzdálenostní metrikou pro shlukování map je právě Kosinova vzdálenost. Ty byla použita i pro finální shlukování sloužící pro porovnání uživatelského shlukování a vyhodnocení výsledků. Na tomto datasetu bylo také zjišťováno, která metoda shlukování z Orange poskytuje lepších výsledků vzhledem ke shlukování podle výtvarného stylu map. Rozdíly ve vytvořených shlucích jsou patrné. Došlo k výraznému promíchání jednotlivých map ve shlucích, nicméně ani v jednom případě nedošlo k jasné diferenciaci podle území nebo podle zdroje, z čehož lze usuzovat, že embedder Painters a některá z metod shlukování nejsou schopny provést shlukování map na základě výtvarného stylu.

8 ANALÝZA VÝTVARNÉHO STYLU TEKTONICKÝCH MAP

Pro analýzu výtvarného stylu tektonických map bylo staženo celkem 25 různých map této tematiky z internetu. Zdroje jednotlivých map jsou v příloze 7. Cílem této analýzy bylo zjistit, jak spolu koreluje grafická náplň jednotlivých map a shlukování těchto map.

Pro zjištění hodnoty náplně mapy bylo využito nástroje pro zjištění grafické náplně map (**GMLMT** – Graphic Map Load Measuring Tool) ve verzi 1.2 vytvořeného Radkem Barvířem (Barvíř, 2021). Tento nástroj pracuje s metrikou založenou na detekci hran pomocí Sobelova filtru. Ten dokáže zachytit jak ostré hrany, tak pozvolné přechody barev, které se v mapách vyskytují. Výstup nástroje je procentuální grafická náplň mapy s rozsahem hodnot od 0 % pro zcela prázdnou mapu do 100 % pro zcela naplněnou mapu. Jako optimální grafická náplň mapy je považována náplň do 30 % (Voženílek, Kaňok, 2011).

Po stažení skriptu nástroje GMLMT je pro správné fungování potřeba tento skript správně umístit do souborového adresáře open-source grafického editoru GIMP. Po úspěšném vložení do adresáře a spuštění GIMPu se nástroj objeví v nabídce *Filtry – Detekce hran*. Před analýzou grafické náplně mapy je potřeba mapu nejprve upravit tak, aby její DPI (dots per inch) bylo 100. Tento krok byl proveden v softwaru Zoner Photo Studio 14. Výsledná grafická náplň jednotlivých map datasetu TURIST je v tabulce 8.1.

Tab. 8.1 Mapy datasetu TURIST ve čtyřech shlucích se svoji grafickou náplní.

shluk 1		shluk 2		shluk 3		shluk 4		
	mapa		mapa		mapa		mapa	
grafická náplň mapy v procentech	8,4	15	17,7	10	6,7	21	11,6	9
	13,6	1	17,9	16	11,9	25	12,1	8
	20	20	18	13	13,5	23	16,3	18
	24,4	3	20,8	12	15	11	19,2	19
			22,5	7	15	14		
			22,5	17	17,1	5		
					17,5	4,0		
					17,9	24		
					20,9	22		
					23,9	6		
				25,7	2			

Data byla rozdělena nehierarchickým shlukováním do čtyř shluků. Využita byla metoda k-Means ++ a Kosinova vzdálenost. Po konzultaci problematiky s docentkou Fišerovou z katedry matematické analýzy a aplikací matematiky Univerzity Palackého v Olomouci bylo provedeno vyhodnocení vlivu náplně na shlukování v prostředí Microsoft Excel a to pomocí statistické metody **ANOVA s jedním faktorem**. Pod

akronymem ANOVA se skrývá sousloví ANalysis Of VAriance, v češtině analýza rozptylu. Tato metoda matematické statistiky umožňuje ověřit, jestli na hodnotu náhodné veličiny (zařazení do shluku) pro určitý prvek (mapa) má statisticky významný vliv hodnota některého znaku (grafická náplň mapy). Nulová hypotéza zní, že grafická náplň map je ve shlucích stejná. Alternativní hypotéza zní, že se grafická náplň ve shlucích liší a je zde tedy vliv náplně na shluky. V Excelu je nutné nastavit zdrojovou oblast dat a také hladinu významnosti α . Ta byla nastavena na 0,05. Výsledná F hodnota testového kritéria je 0,96, kritická hodnota testového kritéria je 3,1 (obr. 8.1). Hodnota F je tedy menší než kritická hodnota a není možné zamítnout nulovou hypotézu, ale potvrdit ji. Podle ANOVY tedy nedošlo k prokázání, že by se grafická náplň ve shlucích lišila. Grafická náplň mapy **nemá vliv** na shlukování.

ANOVA						
<i>Zdroj variability</i>	<i>SS</i>	<i>Rozdíl</i>	<i>MS</i>	<i>F</i>	<i>Hodnota P</i>	<i>F krit</i>
Mezi výběry	69,74778	3	23,24926	0,964582	0,427867	3,072467
Všechny výběry	506,1618	21	24,10294			
Celkem	575,9096	24				

Obr. 8.1 Výsledky statistické metody ANOVA s jedním faktorem.

Vzhledem ke čtyřem pozorováním v prvním a čtvrtém shluku nicméně není zcela jasné, zda je splněna podmínka normálního rozdělení datasetu, která by měla být splněna, spolu s podmínkou stejné variability, pro správné fungování metody ANOVA. Předchozí výsledek byl proto ještě zkontrolován dalším testováním. Pokud není jisté, zda je jsou splněny výše zmíněné podmínky, je vhodnější použít neparametrický přístup například pomocí **Kruskal-Wallisova testu**. Tento test za předpokladu shody rozptylů (lze ověřit například **Squared ranks test for variances**) testuje nulovou hypotézu "mediány grafické náplně jsou ve shlucích stejné", proti alternativní hypotéze "mediány grafické náplně nejsou ve shlucích stejné". V softwaru RStudio bylo využito funkce `conover.test` pracující s **Conover-Imanovým testem**. Ten obsahuje jak Kruskal-Wallisův test, tak mnohonásobné porovnávání mediánů. Nulová hypotéza pro tento test zní: "mediány grafické náplně se ve shlucích neliší". Hladina významnosti byla nastavena na 0,05. Výsledkem testu je p-hodnota rovna 0,28. P-hodnota je tedy vyšší než hladina významnosti. Podle výsledků Conover-Imanova testu (obr. 8.2) není možné nulovou hypotézu zamítnout. Závěrem lze říct, že na hladině významnosti 0,05 se mediány grafické náplně ve shlucích neliší, a tedy grafická náplň **nemá vliv** na shlukování.

```

> 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))
kruskal-wallis rank sum test

data: x and group
kruskal-wallis chi-squared = 3.8213, df = 3, p-value = 0.28

Comparison of x by group
(No adjustment)

```

Col Mean	1	2	3
2	-1.038483 0.1554		
3	0.145760 0.4428	1.488503 0.0757	
4	0.784552 0.2207	1.897918 0.0358	0.804380 0.2151

```

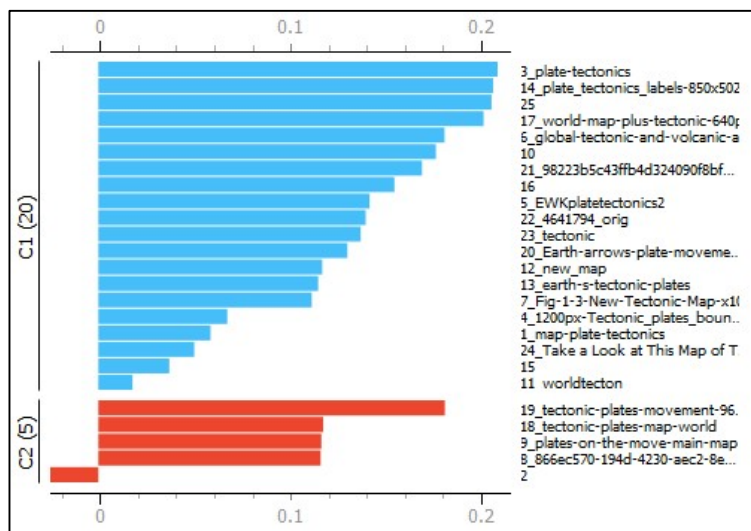
alpha = 0.05
Reject Ho if p <= alpha/2

```

Obr. 8.2 Výsledek Conover-Imanova testu.

8.1 Nehierarchické shlukování datasetu TEKTO

Stejně jako ostatní datasety byl i tento dataset vyhodnocen pomocí nástroje *Silhouette Score* z nehierarchického shlukování (obr. 8.3). Jako ideální počet shluků byly určeny pouze dva shluky (*Silhouette Scores* = 0,101). Hodnota *silhouette score* byla pouze **0,13**. Ze všech datasetů se jedná o nejnižší hodnotu. Způsobené je to pravděpodobně tím, že každá mapa tohoto datasetu TEKTO je z jiného zdroje. V případě datasetu FORECAST a TURIST totiž existují mapy ze stejných zdrojů a u prvního zmíněného dochází ve všech případech ke shlukování podle zdroje. U datasetu TURIST jen v některých případech.



Obr. 8.3 Výsledky nehierarchického shlukování datasetu TEKTO z nástroje *Silhouette Score*.

9 VYUŽITÍ VLASTNÍ NEURONOVÉ SÍTĚ V ORANGE

Při prohledávání dostupných tutoriálů a publikací ohledně tvorby vlastní neuronové sítě v softwaru Orange nebyly nalezeny žádné informace realizovatelnosti takového záměru. Z tohoto důvodu byli kontaktováni vývojáři programu Orange, kteří potvrdili, že tento krok není možný. Možná je tvorba vlastní sítě pouze externě, a to s pomocí modulu Orange SNNS. Lze také natrénovat vlastní model pomocí nástroje *Neural Network* či provést predikci na základě klasifikačních stromů a logistické regrese.

Pro software Orange existuje modul **Orange SNNS** (Stuttgart Neural Network Simulator). Ten umožňuje vytvářet, trénovat a simulovat neuronové sítě ze softwaru SNNS a integrovat je do softwaru Orange. V současné době existují dvě verze modulu OrangeSNNS.py, a to 0.99 a 1.09:

verze 0.99:

- snadněji zobecnitelné pro všechny typy neuronových sítí podporovaných SNNS, protože síťový kód je generován pomocí SNNS,
- rychlejší, protože vyhodnocení sítě je prováděno pomocí programovacího jazyka C namísto Pythonu u verze 1.09,
- alokuje operační paměť.

verze 1.09:

- podporuje dopředné, plně připojitelné vícevrstvé perceptrony se sigmoidální funkcí,
- jednodušší než verze 0.99, protože jsou neuronové sítě vyhodnocovány v Python kódu,
- nealokuje operační paměť (nedochází k tzv. úniku paměti).

V rámci softwaru Orange je pouze možné provést predikci klasifikace (nástroj *Prediction*) na základě klasifikačních stromů a logistické regrese. V tomto případě je nutné všechna zdrojová data, v případě této práce všechny mapy, převést na podobu tabulky. V rámci ní jsou jednotlivými záznamy vybrané mapy a atributy doplňují informace o každé z map.

Další možností je využití nástroje *Neural Network*, který umožňuje provést natrénování modelu na základě vstupních dat, která musí být stejně jako v předchozím případě v podobě tabulky. Tento nástroj pracuje s ANS, které se využívají především na omezené datové sady a na neobrazová data. V tomto případě je problematické převedení dat do tabulkové podoby z důvodu nemožnosti

kvantifikovat barevnost map, tedy jeden z parametrů výtvarného stylu. Barevnost by v důsledku toho nebyla brána v potaz.

Druhou možností je využití již definovaných výtvarných stylů. Ty ale bohužel neexistují tak, jako například existují styly umělecké (impresionismus, realismus). V případě jejich existence by bylo možné pro každou mapu v tabulkovém zobrazení přiřadit atribut konkrétního stylu a na základě těchto dat neuronovou síť naučit. Následně by bylo možné predikovat kategorii, tedy konkrétní výtvarný styl, mapě bez takto přiřazené hodnoty.

Pro navazující práce byla zjištěna řada podnětů. Pokud by byla vyřešena kvantifikace barevnosti map, bylo by možné využít v softwaru Orange možnosti natrénování vlastní neuronové sítě. V případě existence definic jednotlivých výtvarných stylů by bylo dále možné provedení predikce kategorie (výtvarného stylu) map.

10 UŽIVATELSKÉ HODNOCENÍ VÝTVARNÉHO STYLU MAP

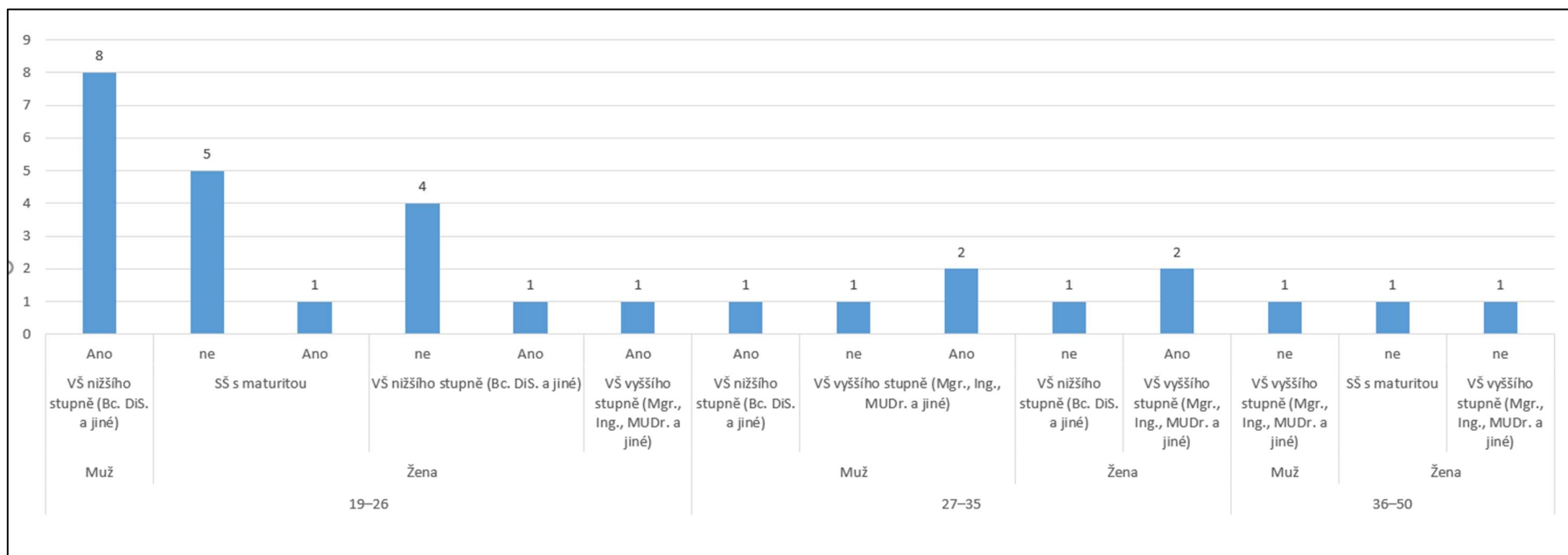
Cílem uživatelského hodnocení tematických map bylo vytvoření shluků ze všech analyzovaných map (25 tematických map v souboru FORECAST, 36 tematických map v souboru TURIST a 25 map v souboru TEKTO) na základě podobnosti výtvarného stylu. Výtvarný styl je definován podle Vondrákové (2020) jako provedení grafiky, kompozičních prvků, jejich parametrů (ukončení linií, jejich tloušťky, výběr fontu...), výběr barev, znakového klíče a provedení layoutu. Rozměry map a poměr jejich stran nebyly měněny, pracovalo se s originálními velikostmi dat.

Uživatelské hodnocení map bylo původně zamýšleno jako šetření v prezenční formě s využitím metod **dotazníku** pro zaznačení obecných informací o testovaných osobách, metody **cart sortingu** a přístupu **thinking aloud**. Cart sorting je metoda, která spočívá v uspořádání souboru prvků (v tomto případě tematických map) do interně homogenních skupin nebo kategorií (shluků) podle určeného principu/kritéria (Spencer, 2009), v tomto případě výtvarného stylu těchto map. V průběhu této fáze by bylo po uživatelích požadováno thinking aloud, tedy shromažďování dat v průběhu testování ze slovního popisu toho, co účastník experimentu zrovna dělá a o čem přemýšlí. Po skončení hlavní části experimentu by byl s testovanou osobou proveden krátký **rozhovor**, v rámci kterého by byl testovaný člověk dotazován na pocity z experimentu, důvody pro zařazení a další otázky doplňující experiment. Během celého experimentu by probíhalo ze strany vedoucího experimentu **direct observation**. Osoba účastnící se experimentu by byla pozorována a vedoucí experimentu by si v průběhu zaznamenával poznatky. Z důvodu pandemie a s tím spojených restrikcí nebylo možné tento experiment podle výše uvedeného postupu provést.

Jako náhradní řešení bylo využito **online dotazníkové šetření** v první polovině března 2021. Nejprve bylo provedeno hierarchické shlukování z embeddovaných snímků embedderem Painters. Nastaveno bylo zvýraznění tří shluků v nástroji *Hierarchical Clustering*. Z každého shluku byl manuálně vybrán jeden typický zástupce, který následně sloužil jako základní mapa shluku, ke kterému měli testované osoby za cíl přiřadit libovolný počet map dané tematiky na základě výtvarného stylu. Pro tyto účely byl vytvořen jednoduchý formulář na platformě Google Forms (<https://forms.gle/WhiXXyyDiebyK8JN8>) (příloha 10). Po aplikaci zpětné vazby z první verze dotazníku byla vytvořena verze druhá, ve které byli uživatelé požádáni o rozřazení tří datových sad tematických map podle výtvarného

stylu do shluků k předem vybraným typickým mapám shluků z hierarchického shlukování. Na konci dotazníku měli testovaní možnost vyplnit základní informace o sobě a také vybrat prvky výtvarného stylu, které upřednostnili během shlukování. Výsledné rozřazení map uživateli bylo následně statisticky vyhodnoceno a porovnáno se shlukováním ze softwaru Orange. Tento přístup experimentu má oproti předchozí prezenční variantě několik nedostatků. Software Orange například porovnává všechny mapy se všemi na základě jejich vzdálenosti podle vybrané metriky. Uživatelé porovnávají jednotlivé mapy pouze se třemi vybranými jako typickými pro daný shluk. Nabízí se alternativní provedení experimentu, kdy by uživatelé měli za cíl pouze rozdělit mapy do shluků bez předem vybrané typické mapy. Díky tomu by bylo dosaženo větší podobnosti s procesem v softwaru Orange, ale také mnohem větších nároků na uživatele. Experiment by byl pro uživatele pracnější a také časově náročnější. Ve všech případech se jedná o sumativní typ experimentu, protože dochází k porovnávání více map stejného tématu mezi sebou.

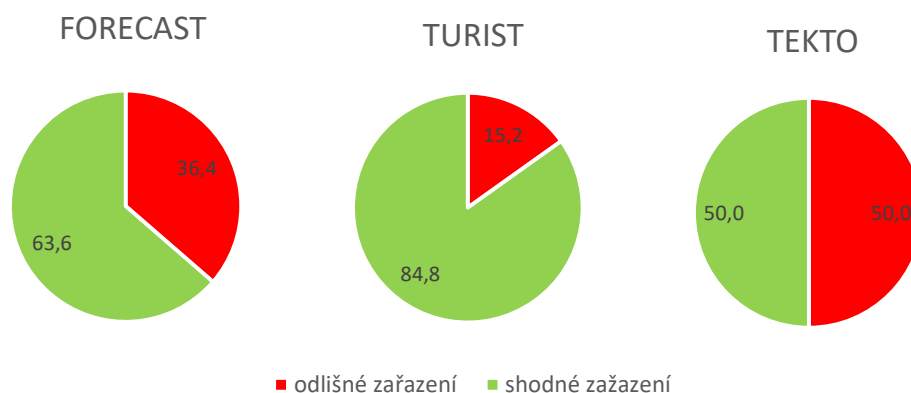
Podle doporučení Nielsena a Pernice (2010) je ideálním počtem respondentů 30, případně 32. Pro tento výzkum se podařilo získat odpovědi právě od **30 osob**. Zastoupení osob podle věku, pohlaví a vzdělání by mělo být v ideálním případě rovnoměrné. To se vzhledem k možnostem distribuce dotazníku a omezenému množství respondentů nepodařilo zajistit. Dotazníkového šetření se zúčastnilo 17 žen a 13 mužů, 20 respondentů bylo ve věku od 19 do 26 let, 7 respondentů od 27 do 35 let a 3 respondenti byly ve věku od 36 do 50 let. Sedm uživatelů uvedlo, že jejich vzdělání je středoškolské s maturitou, 15 respondentů mělo vzdělání nižšího stupně vysoké školy a osm lidí mělo vzdělání vyššího stupně vysoké školy. Dalším zjišťovaným parametrem bylo kartografické vzdělání respondentů. Na základě toho bylo zjištěno, že 16 respondentů má kartografické vzdělání a 14 nikoli. Statistiky počtu respondentů jsou uvedeny na obrázku 10.1.



Obr. 10.1 Počet účastníků dotazníkového šetření podle věku, pohlaví, úrovně vzdělání a kartografického vzdělání.

10.1 Výsledky dotazníkového šetření

Rozřazení map podle výtvarného stylu uživateli bylo statisticky vyhodnoceno v Microsoft Excel. Jednotlivé odpovědi byly sečteny a byla zjištěna procentuální příslušnost dané mapy k jednotlivým shlukům. Následně byl nejčastěji zastoupený shluk u každé mapy porovnán se zařazením mapy do shluku z Orange, tedy s výsledkem procesu embeddingu a hierarchického shlukování. Výsledky jsou prezentovány v podobě grafu pro každý dataset (vázané přílohy 4, 5 a 6). U každého grafu je v tabulkové části barevně znázorněno, do kterého shluku byla mapa přiřazena uživateli a softwarem Orange. Tučně a zeleně je vyjádřeno zařazení mapy v případě shody. Pokud ke shodě nedošlo, je tučnou černou vyznačeno přiřazení v Orange a tučnou červenou přiřazení od uživatelů.



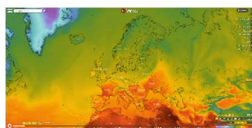
Obr. 10.2 Úspěšnost shodného zařazení map testovaných datasetů do shluků v procentech.

Po agregaci shodného a odlišného zařazení map v jednotlivých datasetech byly výsledky vizualizovány do grafů (obr. 10.2). Nejvyšší úspěšnost byla dosažena u map datasetu TURIST, kdy 84,8 % map je ve shlucích shodných se shluky ze softwaru Orange. Nejhorších výsledků bylo naopak dosaženo u datasetu TEKTO, kdy byla shoda mezi shluky jen v 50 %. Data jsou nicméně v odlišných shlucích zařazována respondenty konzistentně a to s průměrem 70 % (příloha 6). U datasetu FORECAST je shodnost zařazení 63,6 %.

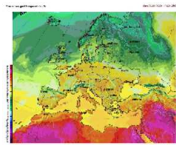
V následujících tabulkách (tab. 10.2, 10.4 a 10.6) jsou znázorněny shody a rozdíly v zařazení map datasetů podle embedderu Painters a následného hierarchického shlukování a podle respondentů. Barevně jsou vyznačeny shluky, kdy jejich barva odpovídá barevnosti shluků v grafech v přílohách 4, 5 a 6. V tabulkách 10.1, 10.3 a 10.5 jsou vybrané typické mapy pro jednotlivé shluky na základě kterých bylo prováděno uživatelské testování.

Tab. 10.1 Vybrané typické mapy pro shluky datasetu FORECAST

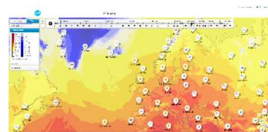
shluk 1



shluk 2



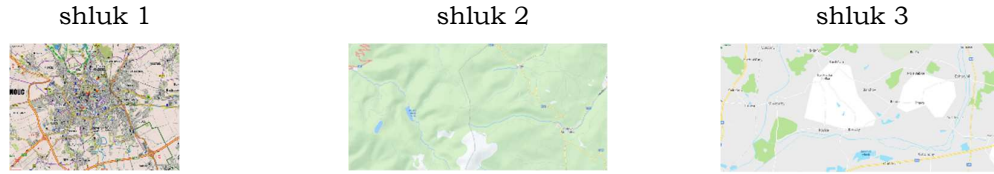
shluk 3



Tab. 10.2 Rozřazení map datasetu FORECAST na základě shod a rozdílů shluků vytvořených embedderem Painters a procesem hierarchického shlukování (HS) a podle uživatelského šetření

	shodné zařazení do shluků	podle Painters a HS	podle respondentů
shluk 1			
shluk 2			
shluk 3			

Tab. 10.3 Vybrané typické mapy pro shluky datasetu TURIST:



Tab. 10.4 Rozřazení map datasetu TURIST na základě shod a rozdílů shluků vytvořených embedderem Painters a procesem HS a podle uživatelského šetření

	shodné zařazení do shluků	podle Painters a HS	podle respondentů
shluk 1	1 2 5 7 10 12 15 18 25 27	4 29	20
shluk 2	13 17 19 31 33	3 9	4 29
shluk 3	6 8 11 14 16 21 22 23 24 26 28 30 32	20	3 9

Tab. 10.5 Vybrané typické mapy pro shluky datasetu TEKTO

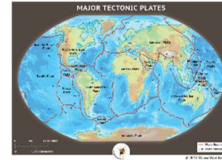
shluk 1













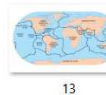




















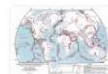

shluk 2



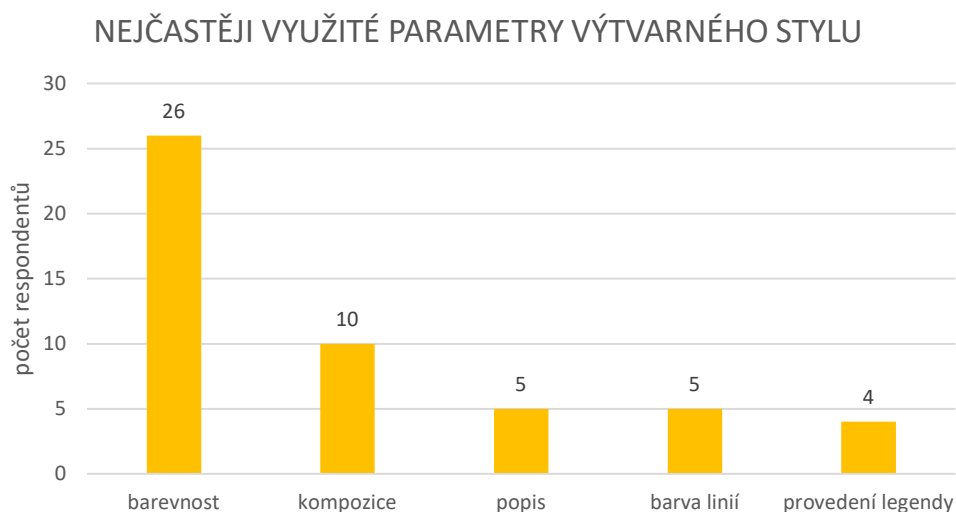
shluk 3



Tab. 10.6 Rozřazení map datasetu TEKTO na základě shod a rozdílů shluků vytvořených embedderem Painters a procesem HS a podle uživatelského šetření

	shodně zařazení do shluků	podle Painters a HS	podle respondentů
shluk 1	 1  18  19  21	 4  5  6  9  11  12  13  15  16  20	 14
shluk 2	 2  7  8  17		 4  6  11  12  13  15  16  20
shluk 3	 3  10  22	 14	 5  9

Z předchozích tabulek 10.1 až 10.6 je zřejmé, že odlišně zařazené mapy jsou ve většině případů přiřazovány ke stejným mapám jako v případě shlukování z Orange. Nicméně došlo k přiřazení k jinému shluku. Výjimku tvoří dataset TEKTO, kde jsou mapy 5 a 9 mimo stejnou skupinu map, která vznikla shlukováním v Orange. Například u datasetu TURIST jsou mapy 4 a 29 umístěny ve shluku 2, zatímco stejné mapy byly v Orange umístěny ke shluku 1. Důvodem bude v tomto případě stínovaný reliéf, který se nachází u vybrané typické mapy pro tento shluk. Respondenti tedy pravděpodobně přiřadily mapy se stínovaným reliéfem k sobě.



Obr. 10.3 Nejčastěji využitě parametry výtvarného stylu při uživatelském testování.

Zajímavé výsledky byly získány od respondentů na otázku, jaké byly nejčastější parametry výtvarného stylu, podle kterých přiřazovaly mapy do shluků. 26 respondentů z 30 uvedlo (86,7 %), že shluky vytvářeli na základě barevnosti. Dále byla zastoupena kompozice, popis, barva linií a provedení legendy. Tyto kategorie uvedlo podstatně méně respondentů. Zařazování map do shluků respondenty probíhalo především podle barevnosti stejně jako v případě shlukování v Orange pomocí embedderu Painters (kapitola 6.2).

Z výsledků je patrné, že byly shluky vytvářeny podle podobného kritéria (barevnost) a u všech datasetů byla shodnost zařazení map do shluků od uživatelů a v prostředí Orange alespoň 50 % (u datasetu TURIST dokonce 84,8 %). Pokud byly v datasetu mapy ze stejných zdrojů, jsou součástí jednoho shluku. Největších rozdílů bylo dosaženo u datasetu TEKTO, kdy je každá mapa z jiného zdroje.

11 VÝSLEDKY

Před samotnou analýzou a vymešováním výtvarných stylů tematických map softwarem Orange bylo nutné získat tři různé datasety těchto map. Dále byly tyto mapy využity pro uživatelský experiment, který sloužil k porovnání výsledků. Konkrétně bylo sesbíráno z internetových zdrojů **25 map předpovědi počasí** (teplota vzduchu) – dataset FORECAST, **36 map turistických/topografických** – dataset TURIST a **25 map tektonických** – dataset TEKTO.

Na datasetech byla provedena **série experimentů**, které měly za cíl zjistit, jak vybraný embedder Painters funguje, jak je ovlivněno shlukování při úpravě parametrů vstupních dat a zda je možné jeho využití pro vymezení výtvarného stylu map. Na datasetu FORECAST byly provedeny experimenty zkoumající **barevnost** (kapitola 6.2), jako jeden z hlavních parametrů výtvarného stylu, a parametry vstupních dat s potenciálem ovlivnit proces shlukování – **ořez, otočení** (kapitola 6.1) a **rozlišení** (kapitola 6.3). Vliv barevnosti na shlukování byl pozorovatelný již při shlukování originálních map (obr. 6.1 a 6.2) i při experimentu s dílčími oblastmi z originálních map (obr. 6.7). Velmi výrazný vliv barevnosti byl prokázán také při experimentu s mapami převedenými do odstínů šedi, kdy došlo k rozdělení vstupního datasetu na dvě poloviny (obr. 6.6). Vliv na shlukování byl prokázán také u rozlišení map a výřezů z originálních map (kapitoly 6.1 a 6.3). Tvorba výřezů se ukázala jako problematická pouze při použití výrazných výřezů (méně než 6 % mapy) z originálních map (obr. 6.3). Různé rozlišení map v důsledku rozdílné velikosti originálních map sice může způsobit odlišné zařazení dané mapy do shluku (obr. 6.8 a 6.9), pokud ale není mapa zmenšena schválně a výrazně (obr. 6.10), není tento fakt problém pro tvorbu výsledných shluků. Podle oficiálních informací autora embedderu by pro jeho správné fungování měl být normalizován poměr stran a rozlišení map, jak je uvedeno v kapitole 6.3. Tímto krokem ale dojde ke ztrátě některých informací v mapě, protože dojde ke zmenšení map a zároveň k jejich ořezu. Méně výrazné prvky mapy v důsledku toho mohou být výrazně potlačeny. Posledním zkoumaným parametrem v této části byla orientace mapy. **Vliv orientace** map na shlukování **nebyl zjištěn**. Naopak bylo z provedených experimentů zjištěno, že **nejvýraznější vliv** na shlukování map má jejich **barevnost** (obr. 6.6, obr. 6.7). Dále byl zjištěn **vliv** na shlukování v důsledku odlišného **rozlišení map** a výrazných **výřezů map**.

Na dalším datasetu, konkrétně datasetu TURIST bylo zjišťováno, která **metoda shlukování** poskytuje lepší výsledky (kapitola 7). Výsledkem procesu embeddování je 2 048 vektorů popisující danou mapu. Pro tvorbu shluků podle výtvarného stylu

je následně nutné využít některou z nabízených metod hierarchického nebo nehierarchického shlukování. Rozdíly ve vytvořených shlucích jsou viditelné, došlo k výraznému promíchání jednotlivých map. Nicméně ani v jednom případě nedošlo k jasné diferenciaci podle území nebo podle zdroje, z čehož lze usuzovat, že embedder Painters a některá z metod shlukování nejsou schopny provést shlukování map na základě výtvarného stylu těchto map. Výsledek z nehierarchického shlukování tohoto datasetu byl vyhodnocen také pomocí *silhouette score*, který slouží k vyhodnocení kvality techniky shlukování. Porovnány byly různé inicializační metody a vzdálenostní metriky. Průměrné hodnoty *silhouette score* se pohybovaly v rozmezí od 0,1 do 0,24. Shlukováním tedy nebylo dosaženo jasného vymezení shluků a některé mapy, s hodnotou *silhouette score* blízkí se nule, jsou na hranici mezi dvěma shluky. Výsledky nejsou příliš jednoznačné. Ani u jedné metod shlukování nedošlo k tvorbě shluků podle výtvarného stylu map a proto **není možné žádnou shlukovací metodu označit za lépe vyhovující** pro daný účel.

Poslední dataset TEKTO byl využit pro zjištění vlivu **grafické náplně** jednotlivých map na shlukování těchto map (kapitola 8). Bylo předpokládáno, že grafická náplň map bude vzhled výsledných shluků ovlivňovat. Pro potvrzení/vyvrácení hypotézy bylo využito statistické metody ANOVA s jedním faktorem a také Conover-Imanova testu. Výsledkem bylo v obou případech přijetí nulové hypotézy, tedy **grafická náplň** mapy **nemá vliv** na shlukování.

Dalším výsledkem práce je **kartografický popis** (kapitola 4) získaných map. Ten byl proveden pro snazší orientaci v dílčích prvcích jednotlivých mapových děl a pro snazší vyhodnocení výsledků. Mapy byly popsány jak slovně, tak s pomocí tabulky, kdy k jednotlivým základním a nadstavbovým kompozičním prvkům mapy a ke grafickému vyhotovení byly přiřazovány hodnoty podle toho, zda daná mapa obsahuje konkrétní prvek. Součástí tohoto dílčího cíle bylo také provedení **uživatelského šetření** (kapitola 10), v rámci kterého bylo zjišťováno, jak a na základě kterých prvků výtvarného stylu budou rozřazovat mapy tří datasetů uživatelé. Také byla zjišťována a vyhodnocována shoda a rozdílnost ve vytvořených shlucích. Zařazování map do shluků probíhalo jak u respondentů, tak v Orange, pomocí embedderu Painters, především podle barevnosti map. U respondentů byla dále zastoupena kompozice, popis, barva linií a provedení legendy. Největší shody bylo dosaženo u datasetu TURIST, kdy bylo shodně přiřazeno 84,8 % map. Naopak nejvíce rozdílné bylo přiřazování map do shluků u datasetu TEKTO, zde byla shodnost přiřazení map pouze 50 %. I přes rozdílnosti byla data v odlišných shlucích zařazována respondenty poměrně konzistentně a to s průměrem až 70 % příslušnosti mapy do shluku.

Výsledkem šetření je zjištění **shodného** dominantního **principu tvorby shluků** uživateli a embedderem Painters a to především na základě **barevnosti**. Dále pak **vyšší míra shodně přiřazených map** do shluků u map z datasetů z **opakujících se zdrojů**.

Relativně dobrých výsledků shlukování podle výtvarného stylu bylo dosaženo u datasetu FORECAST, kdy došlo ke shlukování map ze stejných zdrojů a potenciál embedderu byl prokázán i při shlukování různých území stejných map. Dataset byl shlukován i pomocí nehierarchických metod s průměrným výsledkem *silhouette score* 0,47. U datasetu TURIST naopak nedošlo k jasnému rozdělení map podle zdrojů (tedy podle výtvarného stylu). Mapy jsou až na výjimky ve shlucích promíchané. Výsledek z *silhouette score* byl také mnohem nižší, nežli u prvního datasetu. Průměrná hodnota byla jen 0,24. Ještě nižší hodnoty, a to 0,13, bylo dosaženo u datasetu TEKTO. Vysvětlením může být fakt, že se jedná o jediný dataset, v rámci kterého je každá mapa z jiného zdroje.

Celkově je možné označit embedder Painters a embeddery softwaru Orange obecně za **nevhodné pro vymezení výtvarného stylu map**. Výsledky jsou velmi ovlivněny barevností. To vzhledem k procesu a účelu, za kterým byl embedder natrénován ukazuje na fakt, že používání embedderů/neuronových sítí na jiná data, než na která byla natrénována může negativně ovlivnit výsledky celého procesu. Embedder byl natrénován na uměleckých dílech a je tedy schopen s určitými omezeními shlukovat data na základě uměleckého stylu, případně na základě kartografického stylu. Pokud je ve zpracovávaném datasetu více map z jednoho zdroje s výraznými barvami, jak tomu bylo u datasetu FORECAST, jsou výsledky uspokojující. Shluky jsou tvořeny mapami ze stejného zdroje a jsou dosaženy i dobré hodnoty *silhouette score*. V případě, že jsou shlukovány mapy z různých zdrojů a mapy s nevýraznými barvami, jsou již mapy ve shlucích promíchané a i *silhouette score* dosahuje mnohem nižších hodnot.

12 DISKUZE

V průběhu řešení cílů práce bylo nutné vyřešit řadu problémů. Jako benefity této skutečnosti se objevily podněty pro navazující práce. Nejvíce zatížila exaktní řešení cílů práce absence definic výtvarného stylu. V rámci práce byla získána pouze slovní definice od Vondrákové (2020), ze které bylo také vycházeno. Termín výtvarný styl nemá ani jednoznačný anglický ekvivalent. Existující slovní spojení „*artistic style*“ je označení jak pro styl umělecký, tak pro styl výtvarný. Neexistuje ani žádná práce, která by definovala různé výtvarné styly map nebo alespoň s tímto termínem více pracovala. V literatuře jsou pouze práce analyzující mapy na základě stylů kartografických či uměleckých.

Problematickým faktorem práce je nevhodnost existujících embedderů softwaru Orange pro tento typ analýzy. Neuronové sítě, které nejsou natrénované na analyzovanou problematiku by totiž s největší pravděpodobností mohly vykazovat chybné výsledky, jak tomu bylo i v případě této práce, a proto by bylo vhodnější vytvořit vlastní neuronovou síť, která by byla natrénovaná přímo na problematiku vymezení výtvarného stylu map. Tento proces by byl ale velmi časově a výkonově náročný a jeho uskutečnění by mohlo být náplní několika diplomových nebo disertačních prací. Využitý embedder Painters, který na základě provedeného experimentu ze všech dostupných embedderů poskytuje nejlepší výsledky, byl natrénován na uměleckých dílech světových malířů a je tedy schopen s určitými omezeními shlukovat mapy na základě uměleckého stylu, nikoliv výtvarného.

Výzvou pro další práce je také kvantitativní vyjádření barevnosti map, které by mohlo být následně použito pro predikci výtvarného stylu map nástrojem *Prediction*, jehož využití je ale podmíněného již definovanými výtvarnými styly.

Výsledky z procesu embeddingu je potřeba hodnotit až po procesu shlukování ať už se jedná o shlukování hierarchické nebo nehierarchické, a to buď pomocí dendrogramu nástroje *Hierarchical Clustering* nebo nástroje *Silhouette Plot*. Hodnocení shluků na základě matice z nástroje *Image Grid* je totiž subjektivní a vliv na vymezení shluků s využitím tohoto nástroje závisí na nastavení počtu řádků a sloupců, tedy na velikosti matice.

Problémem při práci byl také relativní nedostatek tematických map z internetových zdrojů. Například některé mapy s teplotou vzduchu jsou prezentovány často v podobě ikon počasí s příslušnou teplotou. Někdy je teplota znázorněna pouze pomocí tabulky s místem a teplotou v tomto místě nebo je na

internetových stránkách pouze v podobě textu nebo videa. Kvůli malému množství map turistických byly tyto mapy doplněny mapami topografickými.



Obr. 12.1 Mapa předpovědi počasí s vyjádřením počasí pomocí ikon s číslem pro teplotu vzduchu (zdroj:

https://www.dwd.de/EN/weather/weather_world/europa/wetter_europa_node.html).

Proces shlukování map komplikuje také rozlišení jednotlivých map. Podle oficiálního popisu autora embedderu Painters by mělo pro správně fungování dojít k normalizaci vstupních dat a to ořezem na čtvercový formát a snížením rozlišení na 256 x 256 pixelů. Krok normalizace vstupních dat má nicméně za následek ztrátu informace u některých map a snížení vlivu méně výrazných částí map a tím zvýšení vlivu barevného provedení. Druhou možností je mapy nenormalizovat a ponechat jejich rozlišení a poměry stran jako jeden z parametrů definujících výtvarný styl map. Normalizací rozlišení se totiž neprokázalo zlepšení výsledků tohoto embedderu pro shlukování map a případného následného vymezení výtvarného stylu map. Ve výsledcích se nicméně rozdíly objevily. Otázkou zůstává, do jaké míry za to může normalizace dat nebo potlačení méně výrazných prvků map. S rozlišením a především s výřezy map byly provedeny i další experimenty, kdy byl zjištěn vliv až při výřezech o velikosti asi 6 % původních map.

13 ZÁVĚR

Cílem diplomové práce bylo identifikovat výtvarné styly několika sad tematických map pomocí programu Orange využívající techniky machine learning. Řešení hlavního cíle bylo rozděleno do následujících čtyř dílčích postupných cílů.

Prvním dílčím cílem bylo z dostupných zdrojů získat několik sad tematických map. Z internetových zdrojů byly získány mapy turistické (TURIST), mapy předpovědi počasí (FORECAST) a tektonické mapy (TEKTO). Tyto mapy byly vybrány vzhledem k jejich různorodosti jejich barevného a výtvarného provedení. Díky tomu obsahují datasey mapy s co možná největší rozmanitostí, na kterých bylo možné provést řadu experimentů zkoumající vliv parametrů map na proces shlukování a také možnosti softwaru Orange pro tento účel.

Druhým dílčím cílem bylo provedení kartografického popisu získaných sad map. Ten byl proveden jak slovně, tak pomocí tabulek, ve kterých bylo hodnoceno více než 30 prvků map pomocí koeficientů (kapitola 4). Popis byl proveden pro snazší a přehlednou práci s mapami. Součástí druhého dílčího cíle bylo provedení experimentu „map use“ (kapitola 10). Ten byl zaměřen na subjektivní ohodnocení map podle výtvarného stylu jednotlivých sad. Experimentu se zúčastnilo celkem 30 respondentů a jeho výsledky byly porovnány se shluky z Orange ve čtvrtém dílčím kroku. Díky experimentu bylo zjištěno, podle jakých parametrů výtvarného stylu budou tvořit shluky uživatelé map a zda budou mapy do shluků přiřazovány shodně nebo rozdílně vzhledem ke shlukování map embedderem Painters.

Třetím dílčím cílem bylo pomocí softwaru Orange zjistit, jak se mapy budou shlukovat podle výtvarných stylů s využitím již integrovaných neuronových sítí (kapitoly 6, 7 a 8). Mapy byly analyzovány s cílem zjistit vliv parametrů map (rozlišení, ořez, otočení) a prvků výtvarného stylu map (barevnost, grafická náplň map) na proces shlukování pomocí metod machine learning softwaru Orange a pomocí dostupných shlukovacích metod (hierarchické a nehierarchické). Provedení tohoto podcíle bylo časově nejnáročnější a jeho splněním bylo zjištěno, že možnosti programu Orange nejsou vhodné pro shlukování map podle výtvarného stylu. Pro správné fungování by bylo potřeba natrénovat vlastní neuronovou síť přímo pro tento účel. Zjištění z toho dílčího cíle mohou být použita jako podněty pro navazující práce.

Čtvrtým cílem bylo porovnání výsledků z druhého a třetího dílčího cíle, tj. shlukování map podle uživatelů z experimentu vůči výsledným shlukům map ze

softwaru Orange využívajícího neuronové sítě (kapitola 10). Porovnávalo bylo zařazení jednotlivých map do shluků podle výtvarného stylu. Přínosem uživatelského experimentu bylo zjištění podobností/rozdílností v procesu tvorby shluků podle výtvarného stylu uživateli a embedderem Painters.

Všechny dílčí podcíle lze považovat za splněné, přestože nebylo možné provést identifikaci různých výtvarných stylů tematických map, a to z důvodu nevhodnosti metod machine learning softwaru Orange pro tento účel. Práce objasnila fungování Orange v oblasti shlukování map metodami machine learning a přinesla řadu podnětů a otázek pro navazující práce.

SUMMARY

The aim of the diploma thesis was to identify the artistic styles of several sets of thematic maps using the Orange software's machine learning methods. The solution of the main goal was divided into the following four partial goals.

The first partial goal was to obtain several sets of thematic maps from available sources. Tourist maps (TURIST), weather forecast maps (FORECAST) and tectonic maps (TEKTO) were obtained from Internet sources. These maps were selected due to their diversity of colour and artistic design.

The second partial goal was to perform a cartographic description of the obtained sets of maps. This was done both verbally and using tables in which more than 30 map elements were evaluated using coefficients. The description was made for easier and clear work with maps. Part of the second sub-objective was to perform a "map use" experiment. It was focused on the subjective evaluation of maps according to the artistic style. A total of 30 respondents participated in the experiment and its results were compared with clusters from Orange.

The third partial goal was to use Orange software to find out how the maps will be clustered according to artistic styles using already integrated neural networks. The maps were analysed to determine the influence of map parameters (resolution, cropping, rotation) and elements of map artistic style (colour, graphic content of maps) on the clustering process using Orange software ML methods and available clustering methods. The implementation of this sub-goal was the most time-consuming and its fulfilment revealed that the possibilities of the Orange are not suitable for grouping maps according to artistic style. For proper functioning, it would be necessary to train your own neural network directly for this purpose.

The fourth goal was to compare the results from the second and third sub-goals, clustering of maps according to users from the experiment against the resulting clusters of maps from Orange software. The classification of individual maps into clusters according to artistic style was compared.

All sub-objectives can be marked as fulfilled. However, it was not possible to identify different artistic styles of thematic maps because the available machine learning methods of Orange software are not capable of this process. The work clarified the functioning of Orange in the field of clustering maps using machine learning methods, but also brought a number of suggestions and questions for follow-up work

POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE

Knižní a časopisecké zdroje

BAČO, Antonín. *Výtvarné styly a techniky v rámci kartografických produktů pro veřejnost*. Praha, 2009. Bakalářská práce. Univerzita Karlova, Přírodovědecká fakulta, Katedra aplikované geoinformatiky a kartografie. Vedoucí práce PhDr. RNDr. Jan Daniel Bláha, Ph.D.

BERKA, Petr. *Dobývání znalostí z databází*. Praha: Academia, 2003. ISBN 80-200-1062-9.

DEAN, Jared. *Big Data, Data Mining, and Machine Learning* [online]. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2014. ISBN 9781118691786. Dostupné z: doi:10.1002/9781118691786.

DEMŠAR, Janez. et al. (2013): *Orange: data mining toolbox in Python*. *Journal of Machine Learning Research*, 14 (1), s. 2349–2353.

DOBEŠOVÁ, Zdena (2019) The Similarity of European Cities Based on Image Analysis. In: Silhavy R., Silhavy P., Prokopova Z. (eds) *Intelligent Systems Applications in Software Engineering*. CoMeSySo 2019 2019. *Advances in Intelligent Systems and Computing*, vol 1046. Springer, Cham. https://doi.org/10.1007/978-3-030-30329-7_31.

GABRYŠ, Martin. *Optimalizace turistických map pomocí eye-tracking testování* [online]. Olomouc, 2019 [cit. 2020-03-21]. Dostupné z: <<https://theses.cz/id/h8tfnu/>>. Diplomová práce. Univerzita Palackého v Olomouci, Přírodovědecká fakulta. Vedoucí práce Mgr. Stanislav Popelka, Ph.D.

KLADIVO, Petr. *Základy statistiky*. Olomouc: Univerzita Palackého v Olomouci, 2013. ISBN 978-80-244-3841-2.

KOSIBA, Matěj. *Klasifikace galaxií pomocí strojového učení* [online]. Brno, 2017 [cit. 2020-03-23]. Dostupné z: <<https://theses.cz/id/qmdky7/>>. Bakalářská práce. Masarykova univerzita, Přírodovědecká fakulta. Vedoucí práce doc. Mgr. Norbert Werner, Ph.D.

LINHART, Jiří; PETRUSEK, Miloslav; VODÁKOVÁ, Alena; MAŘÍKOVÁ, Hana. *Velký sociologický slovník*. Praha: Karolinum. 1996. ISBN 80-7184-310-5.

MOHRI, Mehryar; AFSHIN, Rostamizadeh; AMEET, Talwalkar. *Foundations of Machine Learning*. B.m.: MIT Press. 2012. ISBN 9780262018258.

MURPHY, Kevin P. *Machine Learning: A Probabilistic Perspective*. The MIT Press, 2012. ISBN 0262018020.

NIELSEN, Jakob; PERNICE, Kara. *Eyetracking web usability*. Berkeley: New Riders, c2010. ISBN 978-0-321-49836-6.

O'CONNOR, Thomas; ESKEY, Michael *Scales and Indexes*, Types of, Editor(s): Kimberly Kempf-Leonard, *Encyclopedia of Social Measurement*, Elsevier, 2005, Pages 443-453, ISBN 9780123693983, <https://doi.org/10.1016/B0-12-369398-5/00434-5>.

PRAVDA, Ján. *Mapový jazyk*. 1. vyd. Bratislava: Univerzita Komenského v Bratislavě, 1997. 88 s. ISBN 80-223-1102-2.

PRAVDA, Ján. *Mapový jazyk*. 2. dopl. vyd. Bratislava: Vydavateľstvo UK, 2003. ISBN 8022318094.

RIBEIRO, Daniel; CAQUARD, Sébastien. *Cartography and Art*. The Geographic Information Science & Technology Body of Knowledge (1st Quarter 2018 Edition), John P. Wilson (ed). 2018, DOI: 10.22224/gistbok/2018.1.41.4.

ŘEZANKOVÁ, Hana; HÚSEK, Dušan; SNÁŠEL, Václav. *Shluková analýza dat*. Praha: Professional Publishing, 2007. ISBN 978-80-86946-26-9.

SEDLÁK, Jan. *Rozpoznávání vozidel pomocí metod strojového učení* [online]. Brno, 2016 [cit. 2020-08-04]. Dostupné z: <<https://theses.cz/id/7fdqmx/>>. Diplomová práce. Masarykova univerzita, Fakulta informatiky. Vedoucí práce doc. RNDr. Tomáš Brázdil, Ph.D.

SPENCER, Donna; GARRETT, Jesse James. *Card sorting: designing usable categories*. Brooklyn, N.Y.: Rosenfeld Media, c2009. ISBN 1933820020.

ŠARMANOVÁ, Jana. *Metody analýzy dat*, učební text, VŠB-TU, Ostrava, 2012. ISBN 978-80-248-2565-6, 170 s.

VOŽENÍLEK, Vít. *Aplikovaná kartografie I.: tematické mapy*. 2. vyd. Olomouc: Univerzita Palackého v Olomouci, 2004. ISBN 80-244-0270-x.

Internetové zdroje – bez autora

URL1: *What is Machine Learning? A definition – Expert System*[online]. 2017 [cit. 2020-03-23]. Dostupné z: <https://expertsystem.com/machine-learning-definition/>

URL2: *Orange Data Mining: Image Embedding. Orange Data Mining* [online]. Slovinsko [cit. 2020-04-12]. Dostupné z: <https://orange.biolab.si/widget-catalog/image-analytics/imageembedding/>.

URL3: *ImageNet* [online]. Stanford Vision Lab, Stanford University, Princeton University, 2016 [cit. 2020-04-12]. Dostupné z: <http://www.image-net.org/>

URL4: Distances [online]. Orange Data Mining, 2015 [cit. 2021-01-19]. Dostupné z: <https://orange3.readthedocs.io/en/3.5.0/widgets/unsupervised/distances.html>.

URL5: Home – Keras Documentation [online]. [cit. 2020-04-25]. Dostupné z: <https://keras.io/>

URL6: TensorFlow [online]. [cit. 2020-04-25]. Dostupné z: <https://www.tensorflow.org/>

URL7: Orange SNNS module. Ax5 [online]. [cit. 2020-07-03]. Dostupné z: <http://ax5.com/antonio/orangesnns/>

URL8: The Microsoft Cognitive Toolkit. *Oficiální domovská stránka Microsoft* [online]. © Microsoft, 2020 [cit. 2020-11-13]. Dostupné z: <https://docs.microsoft.com/en-us/cognitive-toolkit/>

URL9: Orange (software). In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2020-11-13]. Dostupné z: [https://en.wikipedia.org/wiki/Orange_\(software\)](https://en.wikipedia.org/wiki/Orange_(software))

URL10: OpenFace [online]. Carnegie Mellon University: Brandon Amos [cit. 2020-11-13]. Dostupné z: <https://cmusatyalab.github.io/openface/>

URL11: GitHub - inecj/painters: Winning solution for the Painter by Numbers competition on Kaggle. *GitHub: Where the world builds software · GitHub* [online]. GitHub, 2020 [cit. 2020-12-16]. Dostupné z: <https://github.com/inecj/painters>

- URL12: Hierarchical Clustering [online]. Orange Data Mining, 2015 [cit. 2021-01-31]. Dostupné z: <https://orange3.readthedocs.io/projects/orange-visual-programming/en/latest/widgets/unsupervised/hierarchicalclustering.html>.
- URL13: *ANN and CNN: Analyzing Differences and Similarities*. Viso Suite [online]. viso.ai, 2021 [cit. 2021-02-04]. Dostupné z: <https://viso.ai/deep-learning/ann-and-cnn-analyzing-differences-and-similarities/>.
- URL14: *Neural network* [online]. Orange Data Mining, 2015 [cit. 2021-02-04]. Dostupné z: <https://orange3.readthedocs.io/projects/orange-visual-programming/en/latest/widgets/model/neuralnetwork.html>.
- URL15: *inejc (Nejc Ilenic)* · GitHub [online]. GitHub, 2021 [cit. 2021-02-12]. Dostupné z: <https://github.com/inejc>.
- URL16: Visual Thesaurus: Artistic style. (2020) *Think map Visual Thesaurus - An online thesaurus and dictionary of over 145,000 words that you explore using an interactive map*. [online]. Thinkmap,. Dostupné z: <https://www.visualthesaurus.com/app/view?word=artistic%20style&lang=en>.
- URL17: *Deep Learning Spreads*. Semiconductor Engineering - Deep Insights For Chip Engineers [online]. SMG, 2021 [cit. 2021-02-16]. Dostupné z: <https://semiengineering.com/deep-learning-spreads/>.
- URL18: *Deep learning vs machine learning*. Customer experience, support and sales blog | Zendesk [online]. Zendesk, 2021 [cit. 2021-02-16]. Dostupné z: <https://www.zendesk.com/blog/machine-learning-and-deep-learning/>.
- URL19: Strojové učení a prostředí MATLAB | HUMUSOFT Blog. *Technické výpočty, řídicí technika, simulace* | Humusoft [online]. HUMUSOFT, 2021 [cit. 2021-02-17]. Dostupné z: <https://www.humusoft.cz/blog/20180817-strojove-uceni/>.
- URL20: *k-Means* [online]. Orange Data Mining, 2015 [cit. 2021-02-18]. Dostupné z <https://orange3.readthedocs.io/en/3.4.0/widgets/unsupervised/kmeansclustering.html>.

Internetové zdroje – s autorem

- BARVÍŘ, Radek. Graphic Map Load Measuring Tool (GMLMT). In: Náplň mapy [online]. Olomouc: katedra geoinformatiky UPOL, 2021 [cit. 2021-03-11]. Dostupné z: http://radiat.cz/wp-content/uploads/2021/01/gmlmt1-2_manual.pdf.
- BECONYTE, Giedre; VILIUVIENE, Rita. *The concept and importance of style in cartography*. Geodesy and Cartography [online]. DOI:10.3846/1392-1541.2009.35.82-91. 2009, 35(3), 82-91 [cit. 2020-11-27]. ISSN 2029-6991. Dostupné z: <https://www.tandfonline.com/doi/abs/10.3846/1392-1541.2009.35.82-91>.
- BERKA, Petr. *Dobývání znalostí z databází*. Praha: Academia, 2003, s. 15. ISBN 80-200-1062-9. Dostupné z: <http://www.digitalniknihovna.cz/mzk/uuid/uuid:7e218600-062d-11e6-a611-005056827e51>.
- BHARDWAJ, Ashutosh. Silhouette Coefficient. *Towards Data Science* [online]. Medium [cit. 2021-03-11]. Dostupné z: <https://towardsdatascience.com/silhouette-coefficient-validating-clustering-techniques-e976bb81d10c>.
- DUŠEK, L., KORÍŤÁKOVÁ, E. *Pokročilé metody analýzy dat v neurovědách*. Masarykova Univerzita Brno [cit. 2021-02-04]. Dostupné z: https://is.muni.cz/el/1411/jaro2017/DSAN02/um/53004109/DSAN02_predn_03.pdf.

ESTEVA, Andre; KUPREL, Brett; NOVOA, Roberto; KO, Justin; SWETTER, Susan; BLAU, Helen; THRUN, Sebastian. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* [online]. 2017, 542(7639), 115-118 [cit. 2020-03-23]. DOI: 10.1038/nature21056. ISSN 0028-0836. Dostupné z: <http://www.nature.com/articles/nature21056>.

GATYS, Leon; ECKER, Alexander; BETHGE, Matthias. *Image Style Transfer Using Convolutional Neural Networks*. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [online]. IEEE, 2016, 2016, s. 2414-2423 [cit. 2020-03-23]. DOI: 10.1109/CVPR.2016.265. ISBN 978-1-4673-8851-1. Dostupné z: <http://ieeexplore.ieee.org/document/7780634/>.

GODEC, Primož; PANČUR, Matjaž; ILENIČ, Nejc; et al. *Democratized image analytics by visual programming through integration of deep models and small-scale machine learning*. *Nature Communications* [online]. 2019, 10(1) [cit. 2020-04-12]. DOI: 10.1038/s41467-019-12397-x. ISSN 2041-1723. Dostupné z: <http://www.nature.com/articles/s41467-019-12397-x>.

GOGUI, I.; SATHIESH, Kumar. *Flower species recognition system using convolution neural networks and transfer learning*. 2017. DOI: 1-6. 10.1109/ICSCN.2017.8085675.
KROGH, Anders. *What are artificial neural networks?* *Nature Biotechnology*, 26(2), 195-197, 2008. doi:10.1038/nbt1386.

MOHANTY, Sharada; HUGHES, David; SALATHÉ, Marcel. *Using Deep Learning for Image-Based Plant Disease Detection*. *Frontiers in Plant Science* [online]. 2016, 7 [cit. 2020-03-23]. DOI: 10.3389/fpls.2016.01419. ISSN 1664-462X. Dostupné z: <http://journal.frontiersin.org/article/10.3389/fpls.2016.01419/full>.

ORY, Jérémie; CHRISTOPHE, Sidonie; FABRIKANT, Sara. *Identification of Styles in Topographic Maps* [online]. [cit. 2020-11-27]. DOI:10.5167/uzh-85358. Dostupné z: https://www.researchgate.net/publication/279757918_Identification_of_Styles_in_Topographic_Maps.

ŠTRUPL, Václav. *Komplexní analýza webových stránek*. In: Václavák [online]. 2008 [cit. 2021-04-17]. Dostupné z: http://www.vaclavak.net/files/komplexni_analyza_webovych_stranek.pdf.

VOLNÁ, Eva. *Neuronové sítě 1* [online]. Ostrava, 2008 [cit. 2020-12-16]. Dostupné z: https://web.osu.cz/~Volna/Neuronove_site_skripta.pdf.

YAMASHITA, Rikiya; NISHIO, Mizuho; DO, Richard Kinh Gian; TOGASHI, Kaori. *Convolutional neural networks: an overview and application in radiology*. *Insights Imaging* 9, 611-629 (2018). Dostupné z: <https://doi.org/10.1007/s13244-018-0639-9>.

ZHOU, Xiran; LI, Wenwen; ARUNDEL, Samantha; LIU, Jun. *Deep Convolutional Neural Networks for Map-Type Classification*. *Cornell University* [online]. 2018 [cit. 2019-03-23]. Dostupné z: <https://arxiv.org/abs/1805.10402v1>.

Slovní citace

RNDr. Alena Vondráková, Ph.D., LL.M. - ústní sdělení ze dne 3. března 2020.

PŘÍLOHY

SEZNAM PŘÍLOH

Vázané přílohy

Příloha 1: Kartografický popis map datasetu FORECAST ohodnocením jednotlivých prvků pomocí koeficientů

Příloha 2: Kartografický popis map datasetu TURIST ohodnocením jednotlivých prvků pomocí koeficientů

Příloha 3: Kartografický popis map datasetu TEKTO ohodnocením jednotlivých prvků pomocí koeficientů

Příloha 4: ZASTOUPENÍ MAP DATASETU FORECAST VE SHLUCÍCH na základě uživatelského testování

Příloha 5: ZASTOUPENÍ MAP DATASETU TURIST VE SHLUCÍCH na základě uživatelského testování

Příloha 6: ZASTOUPENÍ MAP DATASETU TEKTO VE SHLUCÍCH na základě uživatelského testování

Příloha 7: Mapy datasetu TEKTO se zdroji podle čísla mapy

Volné přílohy

Příloha 8: Poster

Příloha 9: DVD

Elektronické přílohy

Příloha 10: Dotazník

Popis struktury DVD

Adresář:

- datasety
 - mapy_dotaznik
 - FORECAST
 - SHLUKY
 - TEKTO
 - TURIST
 - VYSTUPY
 - hodnoceni.xlsx
 - mapy_FORECAST
 - 256
 - cb
 - cb_a_barevne
 - otoceni_orez
 - recko_italie
 - rozliseni
 - mapy_TEKTO
 - stat_vyhodnoceni.xlsx
 - zdroje.docx
 - mapy_TURIST
 - 256
 - cele
 - tabulka.mapy.xlsx
- obr_pro_text
- orange
 - vystupy
 - vyt_styl_tem_mapy.ows
- poster
- rstudio
- text_prace
- webova_stranka

Příloha 1: Kartografický popis map datasetu FORECAST ohodnocením jednotlivých prvků pomocí koeficientů

ZDROJ:	AccuWeather	Dark Sky Map	Eldorado weather	CHMI	Meteoblue	Meteocentrale	Meteoex	Ventusky	Weather online	Windy	yr.no
Základní kompoziční prvky:											
legenda	1	0	1	1	1	1	1	1	1	1	1
měřítka číselné	0	0	0	0	0	0	0	0	0	0	0
měřítka grafické	0	0	0	0	0	0	0	0	0	0	0
tiráž	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0	0,5	0	1
název: nadpis	1	0	1	1	1	1	1	1	1	1	1
podnadpis	0	0	0	0	0	0	0	0	0	0	0
mapové pole	1	1	1	1	1	1	1	1	1	1	1
SOUČET_ZKP:	3,5	1,5	3,5	3,5	3,5	3,5	3,5	3	3,5	3	4
Nadstavbové kompoziční prvky:											
směrová růžice	0	0	0	0	0	0	0	0	0	0	0
zeměpisná síť	0	0	0	0	1	0	0	0	0	0	0
loga	1	0	0	1	0	0	0	1	0	1	1
tabulky	0	0	0	0	0	0	0	0	0	0	0
výšeče	0	0	0	0	0	0	0	0	0	0	1
grafy	0	0	0	0	0	0	0	0	0	0	0
vedlejší mapy	0	0	0	0	0	0	0	0	0	0	0
obrázky	0	0	0	0	0	0	0	0	0	0	0
textové pole	0	0	0	0	0	0	0	0	0	0	0
blokdigramy	0	0	0	0	0	0	0	0	0	0	0
citace	0	0	0	0	0	0	0	0	0	0	0
rejstříky a seznamy	0	0	0	0	0	0	0	0	0	0	0
reklamy	0	0	0	0	0	0	0	0	0	0	0
SOUČET_NKP:	1	0	0	1	1	0	0	1	0	1	2
Grafické provedení mapy:											
rám mapy	1	0	1	1	1	1	1	0	1	0	1
popis	0	1	1	1	1	0	0	1	1	1	1
hranice států	1	1	1	0	1	1	1	1	1	1	1
severní Afrika	1	1	1	1	1	1	0,5	1	1	1	1
severní Evropa	1	1	1	1	1	1	1	1	1	1	1
reliéf	0	0	0	0	1	0	0	0	0	0	0
oblačnost	0	0	0	1	0	0	0	0	0	0	0
tlak	0	0	0	0	0	0	1	0	0	0	0
ikony	1	0	0	1	0	0	0	0	0	0	1
SOUČET_GPM:	5	4	5	6	6	4	4,5	4	5	4	6
Ostatní prvky:											
zoom in/zoom out	1	1	0	0	0	0	0	1	0	1	1
statická	0	0	1	1	1	1	1	0	1	0	0
dynamická	1	1	0	0	0	0	0	1	0	1	1
SOUČET_OP:	2	2	1	1	1	1	1	2	1	2	2
ostatní poznámky											
SOUČET:	11,5	7,5	9,5	11,5	11,5	8,5	9	10	9,5	10	14

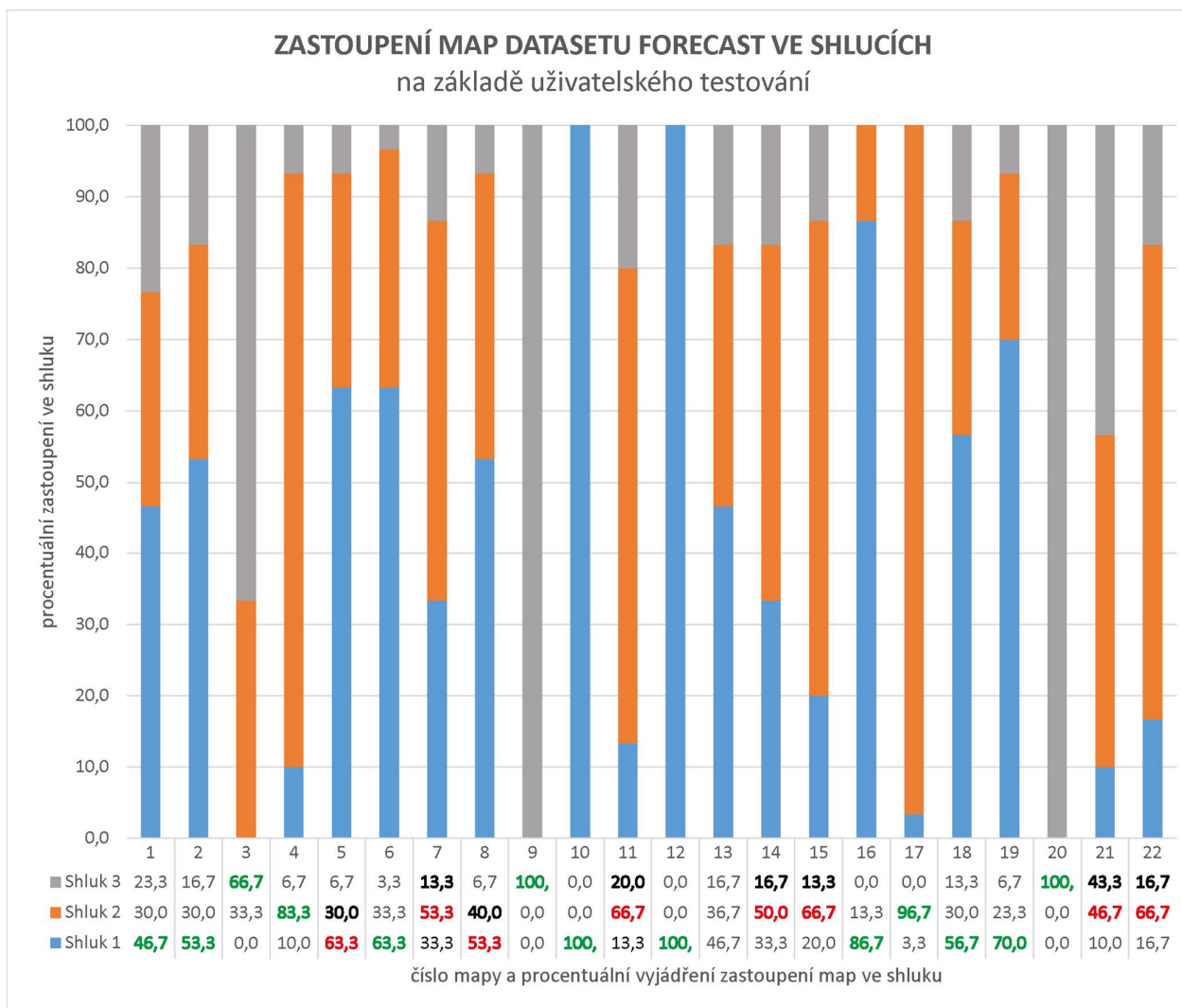
Příloha 2: Kartografický popis map datasetu TURIST ohodnocením jednotlivých prvků pomocí koeficientů

ZDROJ:	bing.com/maps	cykloserver.cz	google.cz/maps	hikebikemap.com	mapy.cz	openstreetmap.org	qwant.com	wego.here.com	yandex.com/maps
Základní kompoziční prvky:									
legenda	0	0	0	0	0	0	0	0	0
měřitko číselné	0	0	0	0	0	0	0	0	0
měřitko grafické	0	0	0	0	0	0	0	0	0
tiráž	0	0	0	0	0	0	0	0	0
název: nadpis	0	0	0	0	0	0	0	0	0
podnadpis	0	0	0	0	0	0	0	0	0
mapové pole	1	1	1	1	1	1	1	1	1
SOUČET_ZKP:	1	1	1	1	1	1	1	1	1
Nadstavbové kompoziční prvky:									
směrová růžice	0	0	0	0	0	0	0	0	0
zeměpisná síť	0	0	0	0	0	0	0	0	0
loga	0	0	0	0	0	0	0	0	0
tabulky	0	0	0	0	0	0	0	0	0
výšeče	0	0	0	0	0	0	0	0	0
grafy	0	0	0	0	0	0	0	0	0
vedlejší mapy	0	0	0	0	0	0	0	0	0
obrázky	0	0	0	0	0	0	0	0	0
textové pole	0	0	0	0	0	0	0	0	0
blokdigramy	0	0	0	0	0	0	0	0	0
citace	0	0	0	0	0	0	0	0	0
rejstříky a seznamy	0	0	0	0	0	0	0	0	0
reklamy	0	0	0	0	0	0	0	0	0
SOUČET_NKP:	0	0	0	0	0	0	0	0	0
Grafické provedení mapy:									
rám mapy	0	0	0	0	0	0	0	0	0
popis	1	1	1	1	1	1	1	1	1
hranice územních celků	1	0	1	1	0	1	1	1	1
reliéf	1	1	1	0	1	0	0	1	0
ikony	0	0	0	0	0	0	0	0	0
SOUČET_GPM:	3	2	3	2	2	2	2	3	2
Ostatní prvky:									
zoom in/zoom out	1	1	1	1	1	1	1	1	1
statická	0	0	0	0	0	0	0	0	0
dynamická	1	1	1	1	1	1	1	1	1
SOUČET_OP:	2	2	2	2	2	2	2	2	2
ostatní poznámky									
SOUČET:	6	5	6	5	5	5	5	6	5

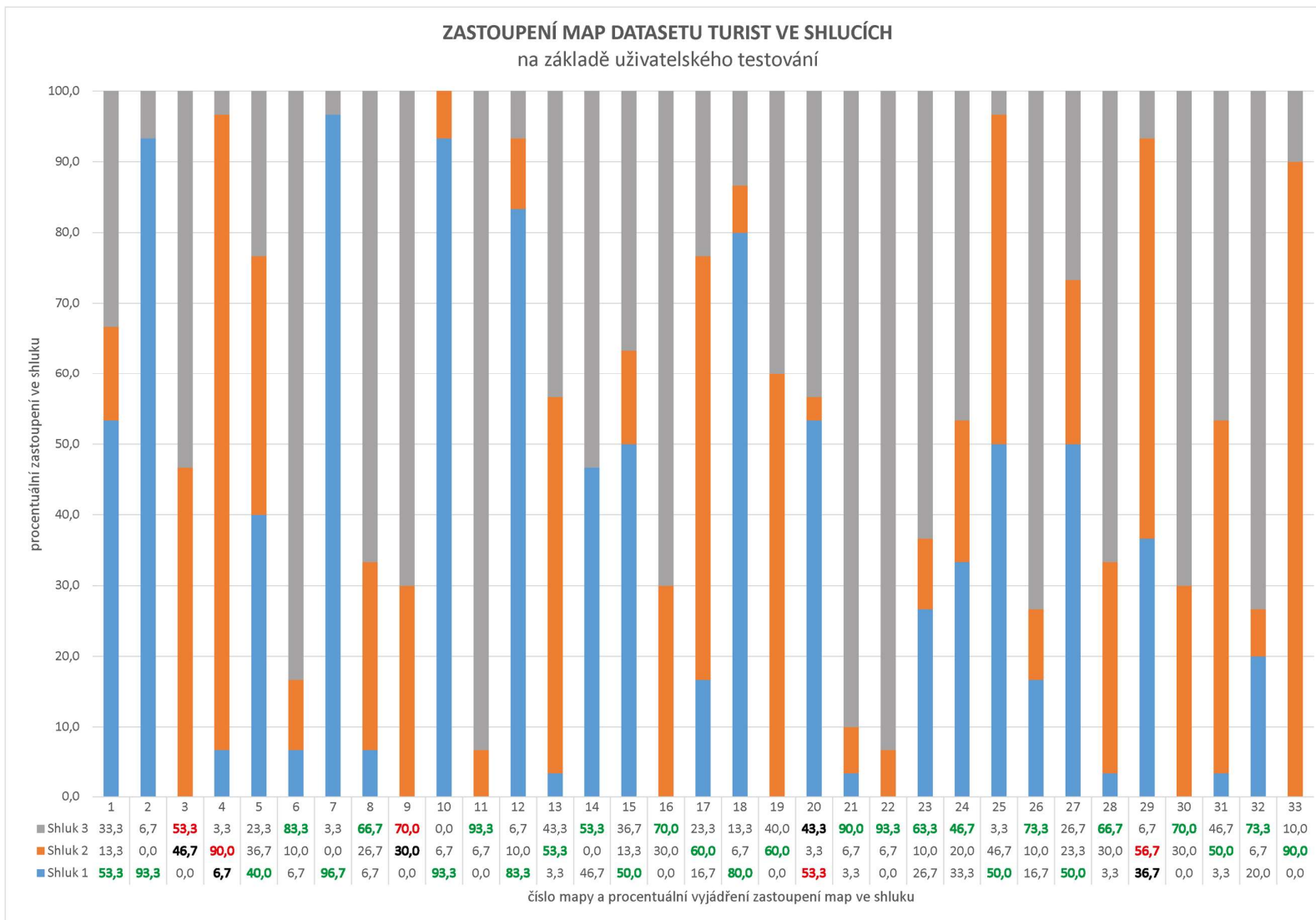
Příloha 3: Kartografický popis map datasetu TEKTO ohodnocením jednotlivých prvků pomocí koeficientů

MAPA ČÍSLO:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Základní kompoziční prvky:																									
legenda	1	0	0	1	1	1	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	0	1	0
měřítka číselné	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
měřítka grafické	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
tiráž	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0,5	0	0	0	0	1	0	0	0	0
název: nadpis	1	1	0	1	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0	1
podnadpis	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
mapové pole	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SOUČET_ZKP:	5	2	1	6	2	4	2	1	1	1	2	1	1	1	3	2,5	1	1	1	1	3	3	1	2	2
Nadstavbové kompoziční prvky:																									
směrová růžice	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
zeměpisná síť	0	0	0,5	1	0	0,5	0	0	0	0,5	1	1	0	0	1	0	0	0	0	0	0	0,5	0	0,5	0
loga	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
tabulky	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
výšeče	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
grafy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
vedlejší mapy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
obrázky	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
textové pole	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
blokdigramy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
citace	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
rejstříky a seznamy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
reklamy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SOUČET_NKP:	1	0	0,5	1	0	0,5	0	0	0	0,5	1	1	0	0	1	0	0	0	0	0	1	0,5	0	0,5	0
Grafické provedení mapy:																									
rám mapy	1	0	0	1	0	1	0	0	0	1	1	1	0	0	0	1	1	0	0	0	1	0	0	1	0
popis	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
hranice států	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Antarktida	1	0,5	0,5	1	0	0,5	0,5	1	1	1	1	0	0	1	1	0	0,5	1	0,5	1	1	0	0,5	0,5	0,5
reliéf	0,5	0	0	0	0	0	0	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
ikony	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SOUČET_GPM:	3,5	1,5	1,5	3	1	2,5	1,5	2	2	3	4	3	1	2	3	2	2,5	2	1,5	2	3	1	1,5	2,5	1,5
Ostatní prvky:																									
zoom in/zoom out	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
statická	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
dynamická	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SOUČET_OP:	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ostatní poznámky																									
SOUČET:	10,5	4,5	4	11	4	8	4,5	4	4	5,5	8	6	3	4	8	5,5	4,5	4	3,5	4	8	5,5	3,5	6	4,5

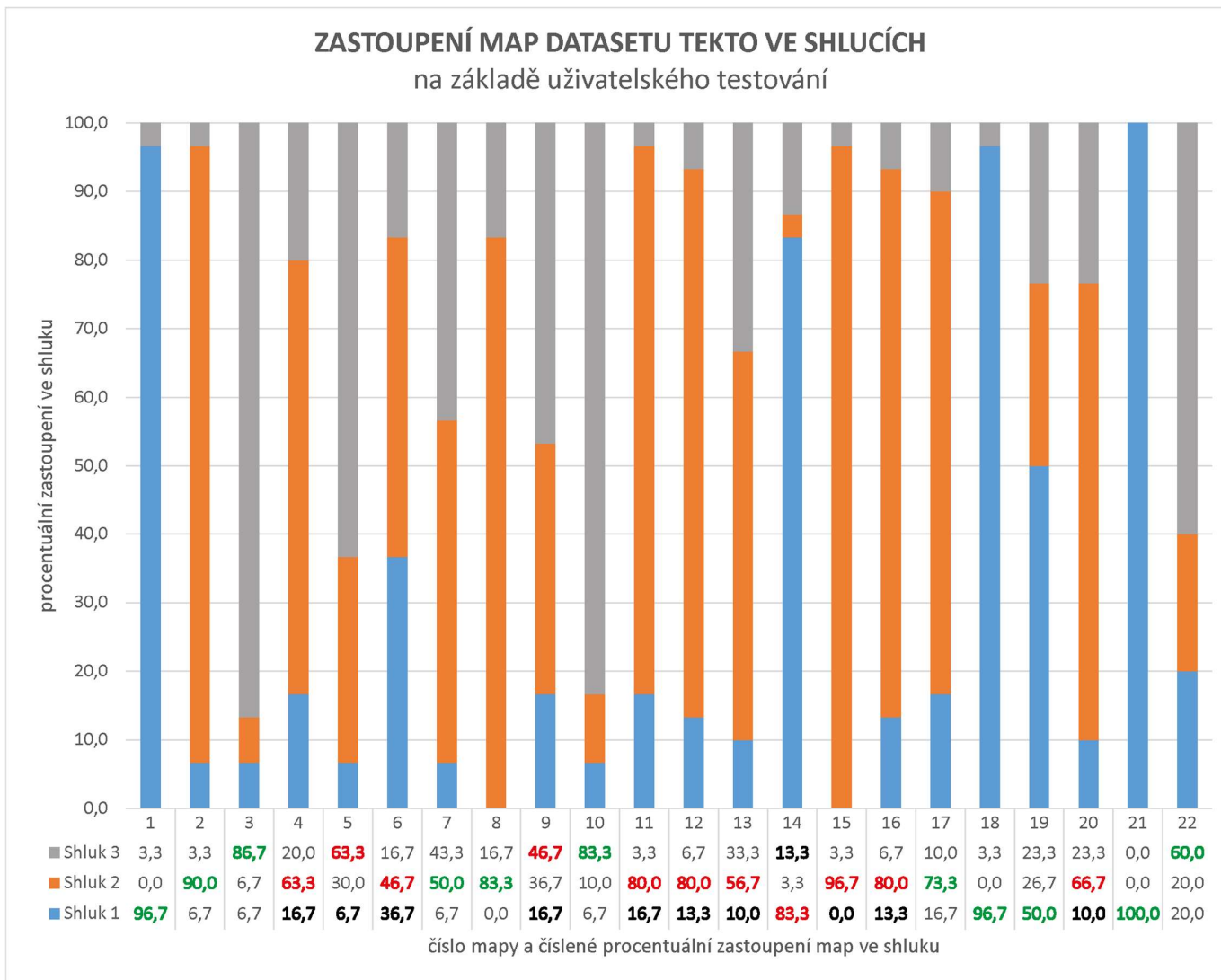
Příloha 4: ZASTOUPENÍ MAP DATASETU FORECAST VE SHLUCÍCH na základě uživatelského testování



Příloha 5: ZASTOUPENÍ MAP DATASETU TURIST VE SHLUCÍCH na základě uživatelského testování



Příloha 6: ZASTOUPENÍ MAP DATASETU TEKTO VE SHLUCÍCH na základě uživatelského testování



Příloha 7: Mapy datasetu TEKTO se zdroji podle čísla mapy

1. <https://www.mapsofworld.com/answers/geography/what-are-plate-tectonics/>
2. <https://www.thoughtco.com/map-of-tectonic-plates-and-their-boundaries-1441098>
3. <https://geology.com/plate-tectonics.shtml>
4. https://commons.wikimedia.org/wiki/File:Tectonic_plates_boundaries_World_map_Wt_180degE_centered-en.svg
5. <https://www.discovermagazine.com/planet-earth/plate-tectonics-the-slow-dance-of-our-planets-crust>
6. <https://geology.com/plate-tectonics.shtml>
7. <https://www.nps.gov/subjects/geology/plate-tectonics-evidence-of-plate-motions.htm>
8. <https://www.nationalgeographic.org/encyclopedia/plate-boundaries/>
9. <https://www.amnh.org/explore/ology/earth/plates-on-the-move2/game>
10. https://www.natur.cuni.cz/fakulta/aktuality/archiv/2011/deskova_tektonika_mapa.jpg/image_view_fullscreen
11. <http://www.johomaps.com/world/worldtecton.html>
12. <https://maheen-geology-portfolio.weebly.com/tectonic-plates.html>
13. <https://www.worldatlas.com/articles/major-tectonic-plates-on-earth.html>
14. <https://earthhow.com/7-major-tectonic-plates/>
15. http://www.merrittcartographic.co.uk/plate_tectonics.html
16. <https://rashidfaridi.com/2009/01/10/plate-tectonics-2/>
17. <https://alsiraatfivesix.wordpress.com/2013/05/19/world-map-with-tectonic-plate-overlay/>
18. https://www.whatarethe7continents.com/what-are-tectonic-plates/?_cfchl_captcha_tk=1fb015a9d9186fdd27b9c91dd72497652ac27720-1613995201-0-AWOGwi3yuvC1fHv1-g6niAKp4HXs_flkWiuFFNip9pbk2Ye4g1t6jhGtMPHcLL4iRF4TQ1BwgR4k3SMEE_OwdZrtep5qXKEWInO44qgZ1k-2EnvKKjE2vw6T1Mgb6_JGRjflbRiv0dLeh0FUMHoN-nxiUY5yCxSZWDoE8fx3uB_JebfVdMvsOSoJKG0tsRghx4Krl0t7FV9v2GRR7rtaJfcSrqRuwEuVULzQ5vM0AgzeTmcWapWXAd402OivEYoahTba1yt8q_CA9pyVBVX9Z4o31rV-hkCzOxOz6RSmjDHqvbFmsYY7aFL99UVh08t3oXnJfAn_X0z0UM2z9RtWaD26O_aeqFSDMNKlrFtYXH0AkV3qJYa-AEewlxR2jzMzI0UrLJPj3OnujJKiY-bgsJaP_9TiL1BkaRFN1Begy3qd08LUWLoITSiOtHzA10-9coewosmXWzcLGzYKaGSO_3NxtpltUBY3bvHRbV7rFWND2_DMIhkKDJTzBd-5qjU2TxSloUj0ny12K9FeG_-RfP6bjj8wdMWb7rE22FHpbhbmUHylRqIDBH2pQpJ3qVnBBhDdg
19. <https://www.bgs.ac.uk/discovering-geology/earth-hazards/earthquakes/what-causes-earthquakes/>
20. <https://www.britannica.com/science/plate-tectonics>
21. <https://www.wildgoose.education/mountains-volcanoes/tectonic-plates-map>
22. <https://geologyswesthead.weebly.com/tectonic-plates--plate-boundaries.html>
23. <https://cosmosmagazine.com/geoscience/how-the-earth-s-tectonic-plates-began-to-move/>
24. <https://cz.pinterest.com/pin/493777546614824530/>
25. <http://planetolog.com/map-world-detail.php?type=TEC&id=4>