

Finální část projektu - varianty projektu

Finální část projektu bude obsahovat primárně tvorbu závěrečné zprávy na základě získaných dat. Projekt se odevzdává 12. ledna 2022, nicméně pamatujte na to, že přes zkouškové období budete mít méně času. Celkově je tato část hodnocena 60 body. Na výběr máte ze dvou variant:

Varianta č. 1: Analýza dopravních nehod

Tato varianta vychází z datasetu, který jste používali v předcházejících částech a je podrobněji specifikovaná v samostatném dokumentu. Obsahuje tři úkoly pokrývající základní témata (tvorba geografických grafů, použití ML modelu (clusteringu), test statistické hypotézy a tvorba závěrečné zprávy).

Varianta č. 2: Analýza zvoleného tématu

Vyberte si zdroj dat z data.bрно.cz dle vlastní preference. Zdrojů může být více a je možné je kombinovat

Tipy:

- [teplotní mapa](#) (v kombinaci např. s [využitím půdy](#) nebo [technickým využitím území](#))
- hlášení [WAZE](#)
- data z cyklistické dopravy ([do práce na kole](#), [aktuální cyklodetektory](#), [historické cyklodetektory](#))
- libovolné další brněnské data

a navrhnete své vlastní zadání (co zhruba chcete z dat získat a prezentovat za informaci), které bude využívat minimálně dva z těchto úkolů

- metodu strojového učení (clustering, regresní model, ...)
- vizualizaci GIS dat
- čištění dat (např. agregace duplicitních hlášení WAZE a podobně)
- statistický test hypotézy

a do 15. prosince 2021 pošlete emailem na pár řádcích, co byste ve svém zadání chtěli dělat na mrizek@fit.vutbr.cz. Po schválení tématu zpracujete a odevzdáte všechny potřebné zdrojové kódy (i pro získání všech používaných údajů), stažená data (ta nemusíte do WISu, ale i mimo WIS, např. do nextcloud FIT) a PDF / HTML soubor s výslednou zprávou. Své řešení budete prezentovat a budete mít možnost detailně vysvětlit své postupy (bude umožněna online forma). **V případě náročnějšího projektu je možné toto zadání dělat ve dvoučlenném týmu (v mailu bude nutné uvést rozdělení zátěže mezi členy).**

Proč se o to pokoušet? Existuje možnost prezentovat získané výsledky magistrátu města Brna, který je může dále využít (s uvedením autora). Vlastní zadání přináší **nevýhodu** v tom, že si sami musíte určit práci jejíž rozsah může být odlišný než u předem daného zadání. Nespornou **výhodou** je však volnost v práci, možnost častějších průběžných konzultací s vyučujícími (záleží na individuální domluvě).

Poznámky k implementaci

Stručnou dokumentaci všech částí (souboru a funkcí) uveďte přímo v odevzdaných souborech. Respektuje konvenci pro formátování kódu PEP 257 [[PEP 257 -- Docstring Conventions](#)] a PEP 8 [[PEP 8 -- Style Guide for Python Code](#)].

Graf by měl splňovat všechny náležitosti, které u grafu očekáváme, tj. měl by být přehledný a jeho velikost by měla být taková, aby se dal čitelně použít v šířce A4 (t.j. cca 18 cm). Toto omezení není úplně striktní, ale negenerujte grafy, které by byly přes celý monitor.

Primárně se počítá, že budete pracovat s následujícími externími knihovnami: *numpy*, *pandas*, *seaborn*, *matplotlib*, *scipy*, *scikit-learn*, *geopandas*, *contextily* a použijete techniky představené na přednáškách. Můžete použít libovolné knihovny zmíněné na přednáškách, další pouze po schválení.

Výsledky budou vyhodnoceny na počítači se systémem Ubuntu 21.10, v prostředí Python 3.9 a s knihovnami v posledních stabilních verzích dle PIP.

Dotazy a připomínky

Dotazy a připomínky směřujte na fórum ve WIS případně na mail mrazek@fit.vutbr.cz.