NASLOV: poletna data science šola v Zvezdarni

NAMEN: naučiti se posredovati spoznanja na podlagi podatkovnih analiz

PROTOKOL:

* **Predstavitev**: Vsak predstavi svoje ugotovitve na izbranih podatkih v obliki vizualizacije, modelov strojnega učenja, … predstavitev je odraz obstoječega znanja predstavljalca in ne zahteva nobenega predznanja. Tudi če kdo naredi zgolj smiselne grafe je to dovolj za predstavitev.
* **Feedback**; primerjava ugotovitev in modelov, vprašanja glede tehnik, s katerimi nismo vsi seznanjeni, vpogledi v kodo, »kako si naredil to in to?« itd.
* **Reševanje problemov**: Sledi konkretno reševanje tehničnih problemov tipa »ne morem v R uvoziti .dta datotek, kako to storim«, »python mi javlja napako, zakaj?« ipd.
* **Nadaljni predlogi**: Vsak na kratko opiše svoj nadaljni predlog podatkov (velikost, domena, atributi in razredi, raziskovalna vprašanja); se posvetujemo in glasujemo za najboljši predlog
* **Bonus**: če se kdo čuti sposobnega, lahko tudi izbere en problem s tega področja in nam ga predstavi. Na primer: različne tehnike obvladovanja manjkajočih vrednosti, algoritme strojnega učenja (na primer random forest), …

PRIPOMOČKI:

* Karkoli potrebuješ za analizo in predstavitev: R, python, Orange, Excel, …

DODATNO:

* Predlogi za skupno opravljanje coursov, …
* **Objava svojih analiz na kaggle platformi**: po lastni odločitvi, če misliš, da je smiselno in koristno.
* Se skupinsko lotimo velikega kaggle tekmovanja
* **Končna refleksija**