

Python projekt podatkovne analitike

Python projekt podatkovne analitike mora zajemati vsaj tri dele (rok 10. april 2022):

- **podatkovni del:** pridobivanje podatkov iz datotek, spletnih virov in/ali API, združevanje podatkov različnih virov, čiščenje in vzorčenje podatkov, predstavitev podatkov s primerno hierarhijo razredov, itd.
- **analitičen del:** npr. osnovna statistična analiza podatkov, univariatna in bivariatna analiza, uporaba netrivialnih algoritmov, metod (ne)nadzorovanega učenja, pristopov analize besedil, slik, omrežij, itd.
- **predstavitveni del:** jasna grafična ali druga predstavitev rezultatov

Predlog projekta mora vsebovati vsaj naslednje postavke (največ pol strani, rok 13. marec 2022):

1. Predlog projekta (t.j. kratek opis namena programa)
2. Glavne funkcionalnosti rešitve (npr. opis primerov uporabe)
3. Terminski plan projekta (t.j. predviden razvoj programa po tednih)
4. Člani razvojne ekipe (t.j. seznam imen ali vpisnih števil študentov)
5. *GitHub* povezava (t.j. spletna povezava do repozitorija projekta)

Končni projekt bo ocenjen na podlagi težavnosti problema, praktične uporabnosti, zajema, priprave in predstavitve podatkov, analize podatkov, primernosti analize, prikaza rezultatov, primernosti prikaza, splošnosti rešitve, delovanja programa in dokumentacije programa.

Spodaj je seznam naslovov (naključno) izbranih projektov iz preteklih let.

- *Primerjava programov in predmetov fakultet*
- *Primerjava demografskih podatkov držav*
- *Analiza Ljubljanskega potniškega prometa*
- *Analiza kaznivih dejanj v Sloveniji*
- *Analiza prometnih nesreč v Sloveniji*
- *Analiza besed v slovenskih slovarjih*
- *Analiza smučarskih skokov skozi čas*
- *Atletika na olimpijskih igrah skozi čas*
- *Analiza rezultatov dirke po Franciji*
- *Analiza ponudbe hrane na e-hrana*
- *Iskanje predlogov podobnih filmov*
- *Napovedovanje dobička filma iz iMDB*
- *Analiza objav na spletnem portalu Reddit*
- *Prikaz Twitter oblakov besed*
- itd.