

Projektni zadatak iz predmeta Intelligentni softverski infrastrukturni sistemi

Potrebno je kreirati aplikaciju za kratkoročnu prognozu potrošnje električne energije na osnovu priloženih podataka. Aplikacija može da sadrži neuronsku mrežu, optimizacione metode, i druge tehnike po želji studenta. Prognoza se vrši za rezoluciju od jednog sata, za vremenski period od jednog do sedam dana.

Aplikacija treba da bude izgrađena u višeslojnoj arhitekturi, pri čemu će se postupci treninga, prognoze, selekcije podataka i cjelokupne poslovne logike izvršavati na servisnom sloju. Poseban sloj su baza podataka, pristup bazi podataka i korisnički interfejs.

APLIKACIJA TREBA DA SADRŽI SLEDEĆE FUNKCIONALNOSTI:

- **UVOZ PODATAKA** - Treba da postoji mogućnost izbora csv datoteka. Nakon izbora datoteka, u bazu podataka se upisuju podaci na satnom nivou za svaki pojedinačan nezavisan i zavisni podatak.
- **TRENING PODATAKA** - Trening podataka se pokreće tako što se izabere određeni datumski opseg, i izabere se opcija „trening“. Nakon izvršenog treninga trenirani modeli se sačuvaju u fajl, a na korisničkom interfejsu se pojavljuje poruka „Trening uspješno izvršen“.
- **PROGNOZA POTROŠNJE** - Prognoza se obavlja tako što se izabere jedan datum, i broj dana za koji se vrši prognoza. Broj dana za koji se vrši prognoza ne može biti veći od sedam. Izborom opcije „Pokreni prognozu“ pokreće se prognoza potrošnje za sate koji se uklapaju u izabrani datum i broj dana za koje se vrši prognoza. Nakon izvršene prognoze, prognozirani podaci se upisuju u bazu podataka i eksportuju se u CSV datoteku. CSV datoteka sadrži kolone:
 - Datum i vrijeme
 - Prognozirano opterećenje

Za jedan datum može se izvršiti više postupaka prognoze. Poslednja izvršena prognoza se uzima kao relevantna.

- **PRIKAZ PROGNOZIRANIH PODATAKA** - Postoji opcija za prikaz prognoziranih vrijednosti prognoze. Prikaz se vrši za izabrani datumski opseg.

NAPOMENE:

- Student dobija podatke u CSV formatu za period od 2018. godine do 2021. godine i to:
 - Istorijske meteorološke podatke
 - Istorijske podatke opterećenja
- Pored postojećih varijabli u dostupnim skupovima podataka, student može dodati dodatne varijable (prosječna potrošnja za prethodni dan, prosječna temperatura za prethodni dan, kalendarski podaci kao što su tip dana, mjesec itd.)

- Kvalitet podataka je presudan faktor o kome ovisi uspješno rudarenje podataka. Vrlo značajnu ulogu u kvalitetu podataka osim izvora podataka imaju i postupci čišćenja i pretprocesiranja podataka. Podaci u izvornom obliku mogu biti nekompletni, atributi mogu imati nedostajuće vrijednosti, ili može postojati nedostatak atributa. Isto tako može se pojaviti nekonzistentnost unutar samih podataka, primjerice nedosljednost u označavanju pojedinih kategorija ili grupa. Neophodno je sve to analizirati i pripremiti podatke prije primjene izabrane metode predviđanja.
- Pored algoritama koji su obrađeni na vježbama, studenti mogu koristiti neki od koncepata optimizacije i mašinskog učenja koji nisu obrađeni na vježbama .
- Prije termina odbrane je potrebno izvršiti treniranje modela i trenirani model sačuvati u fajlu.
- Na odbrani student dobija set nezavisnih podataka za narednih sedam dana koje je potrebno učitati iz csv fajla i upisati u bazu. Na osnovu kreiranog i treniranog modela biće potrebno napraviti prognozu potrošnje električne energije za tih sedam dana. Rezultati se porede sa ostvarenjima za dati period i izračunava se greška (MAPE – prosečna apsolutna procentualna greška).
- Novi set nezavisnih podataka se nalazi u istom obliku kao i trening set podataka.
- Rezultati svakog studenta se rangiraju i oni utiču na konačan broj bodova.
- Boduje se i ispunjavanje arhitekturnih i korisničkih zahtijeva navedenih u tekstu zadatka.