
LV4 –LINEARNA REGRESIJA– ZADACI

Zadatak 1 Predviđanje potrošnje energije uređaja

Cilj zadatka

Primijeniti metode statističke analize podataka na datasetu Predviđanje potrošnje energije uređaja kako biste:

- Proveli deskriptivnu statistiku nad relevantnim varijablama
- Izgradili i analizirali linearni regresijski model
- Procijenili parametre modela, intervale povjerenja i adekvatnost modela
- Izračunati koeficijent korelacije između odabranih varijabli i interpretirati rezultate

Dataset: Korisiti dani [dataset](#) koji sadrži podatke o predviđanjima potrošnje energije uređaja

Koraci za rješavanje zadatka

1. Učitavanju pregled podataka
 - a. Učitati dataset u Python
 - b. Ispitajte i interpretirajte strukturu podataka i provjerite ima li nedostajućih vrijednosti
2. Deskriptivna statistika
 - a. Izračunajte osnovne statistike (apsolutne i relativne srednje vrijednosti s interpretacijama) za odabrane varijable četiri varijable, i interpretirati zašto ste baš njih odabrali.
3. Definiranje hipoteze i odabir varijabli za regresiju:
 - a. Odaberite ciljnu varijablu i jednu nezavisnu varijablu (npr. Appliances (zavisna) , nezavisne: temperature ili relativna vlažnost)
 - b. Pretpostavite linearnu vezu između odabranih varijabli
 - c. Definirati tako za dva regresijka modela
4. Izgradnja linearnog regresijskog modela
 - a. Procijenite parametre modela (nagib i odsječak na osi y)
 - b. Pazite jesu li vam varijable istih mjernih veličina, u slučaju da nisu, morate ih skalirati
 - c. Prikažite dobiven model pomoću scatter-plota s regresijskim pravcem
5. Intervali povjerenja
 - a. Izračunajte 95% intervale povjerenja za parametre regresijskog modela
6. Analiza adekvatnosti modela

- a. Provjerite rezidualnu analizu kako biste procijenili odgovara li linearni model podacima.
 - b. Izračunajte i interpretirajte R^2 kao mjeru kvalitete modela.
7. Izračunajte Pearsonov koeficijent korelacije između ciljane i nezavisne varijable te interpretirajte rezultat
8. Zaključak - napišite zaključak na temelju analize. Jesu li odabrane varijable linearno povezane? Je li model adekvatan za predviđanje?