Umetna inteligenca 2021-2022

Seminarska naloga 1

Janez Sedeljšak, 63200xxx  
Marko Vrečer, 63200329

Vsebina

[1 Vizualizacija podatkov 3](#_Toc89340501)

[2 Ocenjevanje in konstrukcija atributov 4](#_Toc89340502)

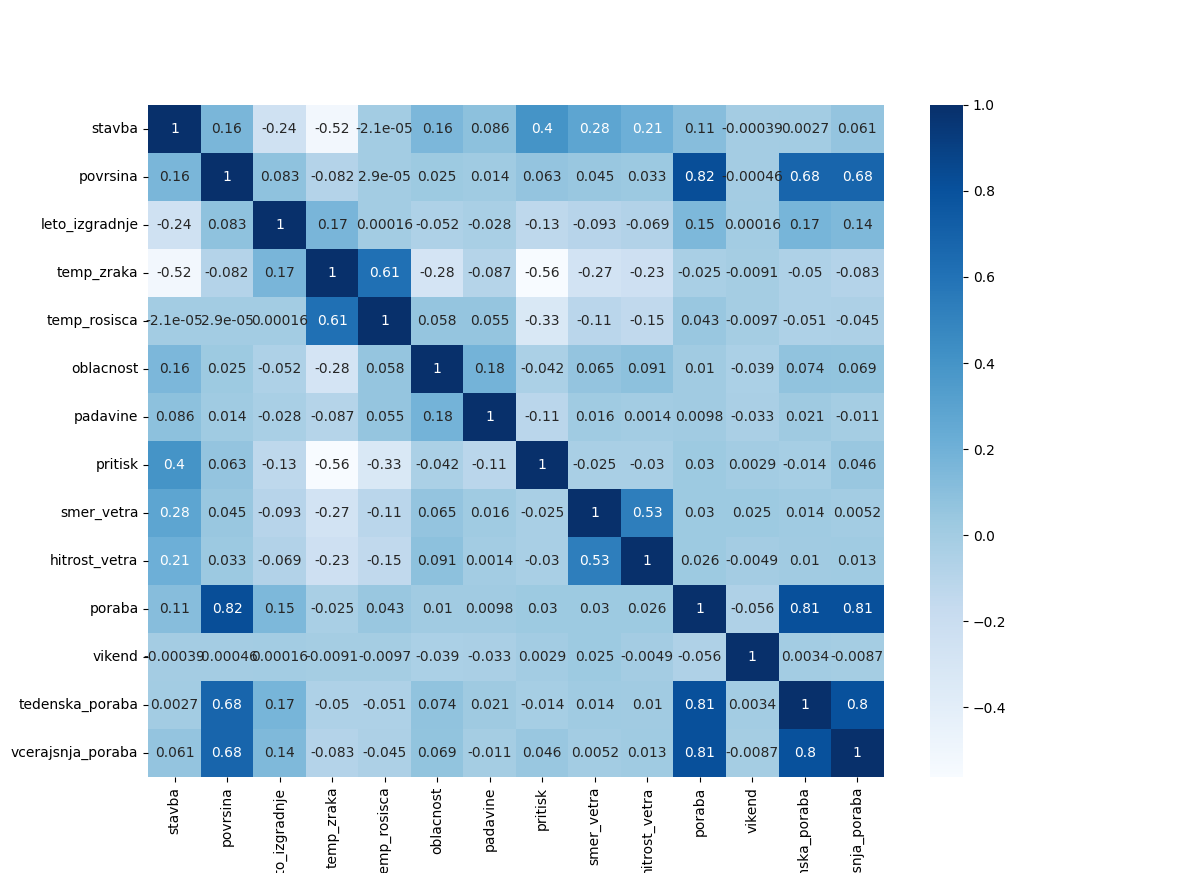
[3 Modeliranje 5](#_Toc89340503)

[4 Evalvacija modelov 6](#_Toc89340504)

# Vizualizacija podatkov

(ostale vizualizacije)

Za lažje razumevanje, kako se atributi med sabo povezujejo, sva naredila skripto v Pythonu, ki pokaže korelacijo med nekaterimi atributi za regresijo. Opazimo lahko na primer močno korelacijo med porabo in tedensko porabo, tedensko\_porabo ter povrsno.



# Ocenjevanje in konstrukcija atributov

## Konstrukcija

(pomoje bi tuki opisal mal ono Parse.java)

## Klasifikacija

## Regresija

Za ocenjevanje atributov pri regresiji sva uporabila knjižnico CORElearn in pognala funkcijo attrEval, ki oceni, koliko je posamezen atribut koristen. Za lažje ugotavljanje vrstnega reda sva rezultat funkcije sortirala, prav tako pa sva za lažje kopiranje združila atribute s presledki.

Uporabila sva več različnih ocen: MSEofMean, RReliefFequalK, RReliefFbestK, MSEofModel in RReliefFsqrDistance, ki so bolj ali manj konsistentno napovedovali, da sta najmočnejša atributa za napovedovanje izpeljana atribut vcerajsnja\_poraba in tedenska\_poraba, kasnejši atributi pa so manj konsistentni: poraba, vikend, leto\_izgradnje...

Rezultati preverjanja so precej pričakovani, saj sva že pri vizualizaciji podatkov opazila, da s porabo predlagani atributi precej korelirajo.

# Modeliranje

## Klasifikacija

## Regresija

### Linearni model – lm

### Regresijsko drevo – rpart

### Support-vector machine – svm

### K-najbližjih sosedov – knn

## Kombinirano učenje

(tuki pomoje kr za oba skupi napiševa)

## Primerjanje uspešnosti glede na podatke iz posameznih regij

# Evalvacija modelov

## Klasifikacija

## Regresija