Matematiniai metodai nekilnojamojo turto vertinime

Doc. Dr. Nomeda Bratčikovienė

Turinys

- Įvadas į duomenų analitiką NT vertės vertinime
  - Vizualizacijos vaidmuo: histogramų, sklaidos diagramų interpretacija
  - Trūkstamų reikšmių nustatymas ir tvarkymas
  - Išskirčių analizė
  - Išskirčių įtaka statistiniams modeliams
- Statistiniai pagrindai
- Regresinė analizė NT rinkoje
- Netiesinė regresija

NT vertinimas – nuo duomenų iki vertės

- Duomenų šaltiniai:
  - Registrų centro duomenys,
  - NT portalai,
  - Savivaldybių duomenys

- Kodėl NT vertintojui svarbu išmanyti duomenų analitiką?
  - duomenų analizė = pagrįstas sprendimas
  - duomenų analitika padeda priimti pagrįstus vertinimo sprendimus, o ne „pakeičia vertintoją“

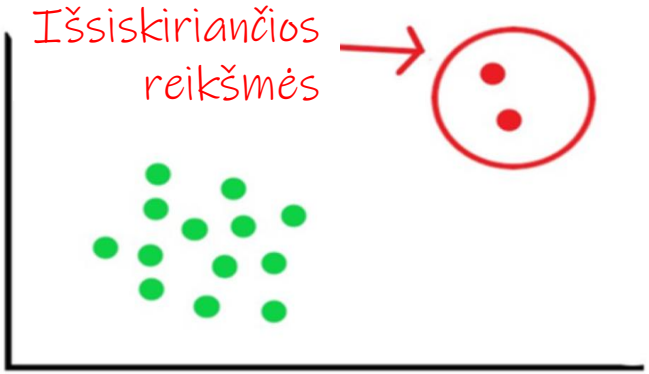
Duomenų kokybė ir praleistos reikšmės

- Jei duomenys *netikslūs* ar *nepilni*, net ir pažangiausi modeliai pateiks klaidingus rezultatus, todėl **duomenų kokybės užtikrinimas** yra pirmas žingsnis į patikimą NT vertinimą.

  - Dažniausios duomenų problemos:
    - praleistos reikšmės,
    - klaidos,
    - išsiskiriantys stebėjimai

Išsiskiriančios reikšmės ir jų aptikimas

- **Išsiskiriančios reikšmės** - reikšmės, kurios yra neįprastos, netipinės arba skiriasi nuo bendrų, standartinių reikšmių.
- Pavyzdžiui NT rinkoje:
  - Tarkime, turime 1000 butų duomenis. Dauguma kainuoja tarp **100 000–250 000 €**, bet vienas objektas – **1 000 000 €**. Šis objektas bus laikomas išskirtimi.



Koreliacinė analizė

Koreliacinė analizė duoda atsakymą į klausimą: ar ryšys tarp kintamųjų egzistuoja ir ar jis reikšmingas?

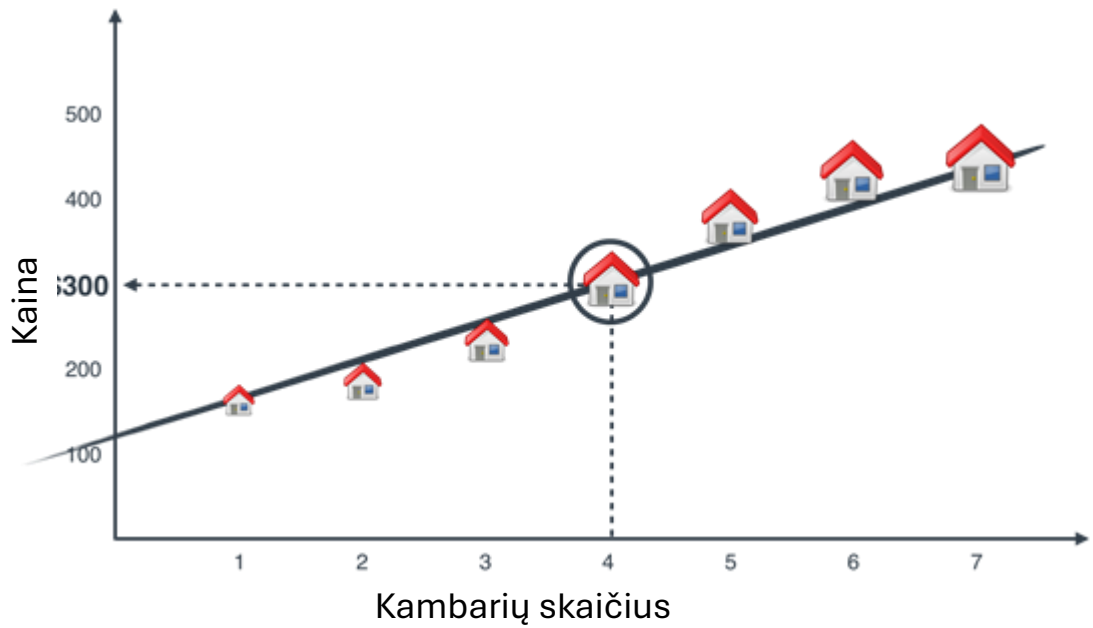
Ryšio matai:

- Porinis koreliacijos koeficientas
- Spearmano koreliacijos koeficientas
- Kendallo τ ranginės koreliacijos koeficientas koeficientas

Tiesinė regresija NT vertinime

- Regresinėje analizėje sprendžiami vieno ar kelių kintamųjų (priklausomų) prognozavimo remiantis kitais (nepriklausomais) uždaviniai.

$$y_i = a_0 + a_1 x_{1i} + a_2 x_{2i} + \ldots + a_k x_{ki} + \varepsilon$$



Vertinimo paklaida

- Prognozės paklaida yra skirtumas tarp faktinės ir įvertintos reikšmių

$$e_i = X_i - \hat{X}_i$$

- Prognozės paklaida gali susidaryti dėl:
  - Neteisingai parinkto modelio
  - Netiksliai įvertintų parametrų, duomenų nepakankamumo
  - Dėl nežinomų ateities paklaidų (baltojo triukšmo) reikšmių
  - Prognozuojamo proceso nepastovumo

Modelių palyginimas

- ME - vidutinė paklaida:

$$ME = \frac{1}{n}\sum_i \left(X_i - \hat{X}_i\right)$$

Adekvačiai sudaryto modelio ME lygi ar labai artima nuliui.

- MAE – vidutinė absoliutinė paklaida:

$$MAE = \frac{1}{n}\sum_i \left|X_i - \hat{X}_i\right|$$

Kai lyginamas faktinės ir teorinės reikšmės atitikimas šis rodiklis vadinamas vidutiniu absoliutiniu nuokrypiu

Modelių palyginimas

- SSE – prognozės likučių kvadratų suma:

$$SSE = \sum_i \left(X_i - \hat{X}_i\right)^2$$

- MSE – vidutinė kvadratinė paklaida:

$$MSE = \frac{1}{n-k} \sum_i \left(X_i - \hat{X}_i\right)^2$$

čia *n* - stebėjimų, *k* – modelio parametrų skaičius.

- RMSE – šaknis iš vidutinės kvadratinės paklaidos:

$$RMSE = \sqrt{\frac{1}{n-k} \sum \left(X_i - \hat{X}_i\right)^2}$$

- AIC, BIC

Prognozių tikslumo rodikliai

- MPE – vidutinė procentinė paklaida:

$$MPE = \frac{1}{n}\sum_i \left(\frac{X_i - \hat{X}_i}{X_i} \right) \cdot 100\%$$

- MAPE – vidutinė absoliutinė procentinė paklaida:

$$MAPE = \frac{1}{n}\sum_i \left| \frac{X_i - \hat{X}_i}{X_i} \right| \cdot 100\%$$