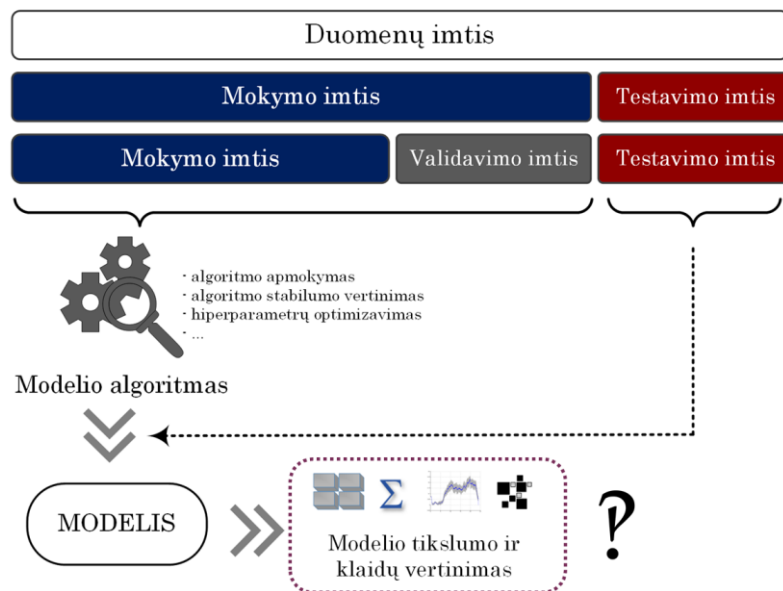


Prižiūrimojo mašininio mokymosi modelio patikra

II dalis. Prižiūrimojo mašininio mokymosi modelio patikra, kai tikslinis kintamasis yra kategorinis (klasifikavimo uždavinys)

Prisiminkime jau matytą modelių patikros principinę schemą (1 pav.), kuri trumpai buvo pristatyta pirmoje šios temos dalyje.



1 pav. Modelio patikros schema

Dabar ypatingą dėmesį skirsime modelio tikslumo ir klaidų vertinimo metrikoms, pagal kurias ieškomas geriausias klasifikavimo modelis.

Klasifikavimo uždaviniuose tikslinis kintamasis yra kategorinis. Dviejų kategorijų pavyzdys – laimėjo-nelaimėjo; jei trys kategorijos, tai, pavyzdžiui, laimėjo-nelaimėjo-lygiosios ir pan. Pačiu paprasčiausiu atveju turime dvi reikšmes arba dvi kategorijas. Gan dažnai vartojamas terminas klasės, ypač kai sudaromas klasifikavimo uždavinys. Paprastumo dėlei, jas pažymėkime 1 arba 0. Įprasta „1“ žymėti teigiamą atsakymą į rūpimą klausimą: ar taps įmonė nemokia? Ar susirgs tam tikra liga? Ar gaus paskolą? Ar laimės loterijoje? Ar išlaikys egzaminą? Anglų kalboje vartojamas atitinkmuo *positive*. Modelio validavimo ar testavimo metu yra žinomos tikrosios reikšmės (klasės), kurias taip pat pažymėkime atitinkamai 1 arba 0. Taigi, atliekant modelio patikrą šiuo atveju iš viso galimos keturios skirtingos kombinacijos: (1; 1), (1; 0), (0; 1), (0; 0), kas atitinkamai žymi (faktas; prognozė).

Gautų keturių kombinacijų interpretacija:

- (1; 1) – modelis teisingai prognozuoja klasę „1“. Tokių atvejų kiekį validavimo / testavimo imtyje žymėkime **TP** (angl. *true positive*);
- (1; 0) – modelis klaidingai prognozuoja klasę „1“, nes prognozavo klasę „0“, nors tikroji klasė yra „1“. Tokių atvejų kiekį validavimo / testavimo imtyje žymėkime **FN** (angl. *false negative*). Statistikos moksle žinoma kaip II rūšies klaida;

Masinis atvirasis internetinis kursas „Dirbtinis intelektas“

- (0; 1) – modelis klaidingai prognozuoja klasę „0“, nes prognozavo klasę „1“, nors tikroji klasė yra „0“. Tokių atvejų kiekį validavimo / testavimo imtyje žymėkime **FP** (angl. *false positive*). Statistikos moksle žinoma kaip I rūšies klaida;
- (0; 0) – modelis teisingai prognozuoja klasę „0“. Tokių atvejų kiekį validavimo / testavimo imtyje žymėkime **TN** (angl. *true negative*).

Panagrinėkime pavyzdį (2 pav.). Tarkime, kad modelio klasifikavimo tikslumo ir klaidų vertinimui yra atidėta 20 stebėjimų, kurių reikšmės yra žinomos ir pateiktos stulpelyje „Faktas“. Tikslinis kintamasis yra kategorinis – 2 klasės: 0 ir 1. Tarkime, kad taikytas tam tikras klasifikavimo modelis. Jo prognozuojamos reikšmės kiekvienam iš šių stebėjimų yra pateiktos stulpelyje „Prognozė“. Taip pat turime dvi klases: 0 ir 1. Lyginant fakto ir prognozės reikšmes eilutėmis, gaunamos visos keturios kombinacijos, kurias, kaip jau buvo sutarta, atitinkamai žymime **TP**, **TN**, **FP**, **FN**.

| ID | Faktas | Prognozė | Išvestis |
|----|--------|----------|-----------|
| 01 | 0 | 0 | <i>TN</i> |
| 02 | 1 | 1 | <i>TP</i> |
| 03 | 1 | 0 | <i>FN</i> |
| 04 | 1 | 1 | <i>TP</i> |
| 05 | 1 | 1 | <i>TP</i> |
| 06 | 1 | 0 | <i>FN</i> |
| 07 | 0 | 0 | <i>TN</i> |
| 08 | 1 | 1 | <i>TP</i> |
| 09 | 0 | 0 | <i>TN</i> |
| 10 | 0 | 1 | <i>FP</i> |

| ID | Faktas | Prognozė | Išvestis |
|----|--------|----------|-----------|
| 11 | 1 | 1 | <i>TP</i> |
| 12 | 0 | 0 | <i>TN</i> |
| 13 | 1 | 1 | <i>TP</i> |
| 14 | 1 | 0 | <i>FN</i> |
| 15 | 0 | 1 | <i>FP</i> |
| 16 | 1 | 1 | <i>TP</i> |
| 17 | 0 | 0 | <i>TN</i> |
| 18 | 0 | 0 | <i>TN</i> |
| 19 | 1 | 1 | <i>TP</i> |
| 20 | 0 | 0 | <i>TN</i> |

2 pav. Pavyzdys

Modelio validavimo / testavimo metu išvesties reikšmės (klasės) yra surašomos į **klaidų (sumaišymo) matricą** (angl. *confusion matrix*), kurios elementai yra kiekvienos kombinacijos atvejų skaičius (3 pav.).

| | | Prognozė | |
|--------|---|-----------|-----------|
| | | 1 | 0 |
| Faktas | 1 | <i>TP</i> | <i>FN</i> |
| | 0 | <i>FP</i> | <i>TN</i> |

3 pav. Klaidų matrica

Pateikto pavyzdžio atveju (2 pav.) yra užpildoma klaidų matrica atitinkamai suskaičiuojant **TP**, **TN**, **FP**, **FN** reikšmių pasikartojimus (4 pav.).

| | | Prognozė | |
|--------|---|----------|---|
| | | 1 | 0 |
| Faktas | 1 | 8 | 3 |
| | 0 | 2 | 7 |

4 pav. Klaidų matrica pateiktam pavyzdžiui

Panagrinėkime dažniausiai naudojamas metrikas, kurios tiesiogiai išskaičiuojamos iš klaidų matricos:

$$\begin{aligned} \textbf{Tikslumas} &= \frac{TP + TN}{TP + FP + FN + TN} \\ (\text{angl. } accuracy) \end{aligned}$$

$$\begin{aligned} \textbf{Preciziškumas} &= \frac{TP}{TP + FP} \\ (\text{angl. } precision) \end{aligned}$$

$$\begin{aligned} \textbf{Išsamumas, jautrumas} &= \frac{TP}{TP + FN} \\ (\text{angl. } recall, sensitivity) \end{aligned}$$

$$\begin{aligned} \textbf{Specifiškumas} &= \frac{TN}{TN + FP} \\ (\text{angl. } specificity) \end{aligned}$$

Šių metrikų reikšmių sritis $[0; 1]$. Siekiama, jog šios reikšmės būtų kuo didesnės. Trumpai pasigilinkime į jų prasmę.

Tikslumas – tai teisingai prognozuotų atvejų iš visos imties skaičius. Ši metrika yra geras pasirinkimas, kai tikslinio kintamojo klasių skaičius „1“ ir „0“ imtyje yra panašus, t. y. mokymo imtis subalansuota. Deja, realaus taikymo uždaviniuose retai taip sutinkama. **Preciziškumas**, **specifiškumas** ir **išsamumas** matuoja modelio klasifikavimo tikslumą ir klaidų tam tikros klasės atžvilgiu. Šios metrikos ypač rekomenduotinos, kai mokymo imtis yra nesubalansuota. Literatūroje yra siūloma ir daugiau metrikų, kurios skirtos vertinti modelio gerumą nesubalansuotoms imtims. **Preciziškumas** rodo, kiek iš prognozuotų „1“ atvejų iš tikrųjų yra „1“. **Išsamumas** rodo, kokia faktinių „1“ atvejų dalis prognozuota teisingai. Beje, terminas **jautrumas** naudojamas kaip sinonimas **išsamumui**. **Specifiškumas** rodo, kokia faktinių „0“ atvejų dalis prognozuota teisingai. Dažnai **preciziškumas** ir **išsamumas** yra įvertinami kartu apskaičiuojant F_1 kriterijų, t. y.

$$F_1 = 2 \frac{\text{preciziškumas} \cdot \text{išsamumas}}{\text{preciziškumas} + \text{išsamumas}}$$

Kuo F_1 reikšmė didesnė, tuo modelis yra geresnis, t. y. pateikia prognozes, kurios maksimaliai sutampa su validavimo / testavimo imtyje esančiais stebėjimais.

Tęskime jau nagrinėtą pavyzdį (2 pav.) ir apskaičiuokime aprašytas metrikas.

$$\text{Tikslumas} = \frac{8 + 7}{8 + 3 + 2 + 7} = 0,75$$

$$\text{Preciziškumas} = \frac{8}{8 + 2} = 0,8$$

$$\text{Išsamumas} = \frac{8}{8 + 3} \approx 0,73$$

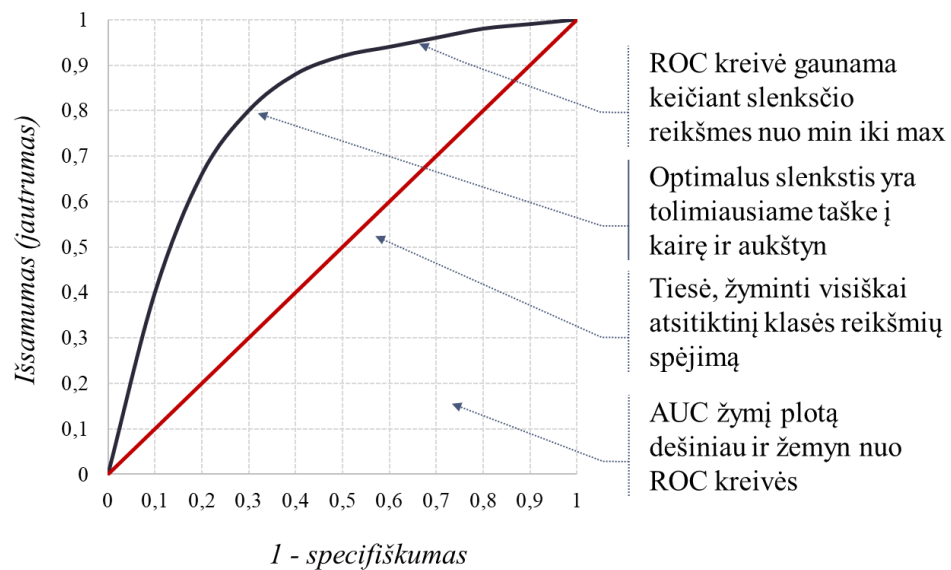
$$\text{Specifiškumas} = \frac{7}{7 + 2} \approx 0,78$$

$$F_1 = 2 \frac{0,8 \cdot 0,73}{0,8 + 0,73} \approx 0,76$$

Visos klaidų metrikos įgyja reikšmes nemažesnes nei 73 proc. Gauti metrikų įverčiai pakankamai neblogi, tačiau kyla klausimas, ar įmanoma šias klasifikavimo prognozes pagerinti parinkus kitą modelį. Todėl visuomet rekomenduojama išbandyti bent keletą modelių ir lyginant pasirinkti tą, kuris yra geresnis. Kryžminės patikros demonstravimas nėra įtrauktas į šį pavyzdį.

Realaus taikymo uždaviniuose, kai modelis taikomas nesubalansuotoms imtims, labai svarbu įsitikinti, kaip kinta modelio paklaidos keičiant **sprendimo priėmimo slenkstį** (angl. *threshold*). Priežastis ta, jog modelis pateikia klasių prognozę „1“ arba „0“ pagal algoritmo apskaičiuotą atitinkamos klasės tikimybę. Pagal nutylėjimą taikomas slenkstis 0,5. Šio slenksčio reikšmės keitimas ir modelio gerumo vertinimas yra analizuojamas braižant **ROC kreivę** (angl. *receiver operating characteristic curve*) ir apskaičiuojant **plotą po šią kreivę AUC** (angl. *area under curve*).

ROC kreivė sudaroma atvaizduojant išsamumo (ROC atveju dažniau yra naudojamas termino sinonimas – *jautrumas*) ir specifiškumo sąryšį (5 pav.). Tiesa, horizontalioje ašyje atvaizduojama (1 – *specifiškumas*).



5 pav. ROC kreivė

Taikant daug modelių, patogų visas ROC kreives atvaizduoti viename brėžinyje ir palyginti tarpusavyje. Tačiau išrinkti geriausią modelį turint daug ROC kreivių būtų sudėtinga, dėl to yra skaičiuojamas AUC. Iš daugelio taikytų modelių rekomenduojama pasirinkti tą, kurio AUC didžiausias.

Apibendrinant pastebėsime tai, jog:

- rekomenduojama išbandyti keletą klasifikavimo modelių. Tuomet apskaičiavus klaidų metrikas, išsirinkti geresnį modelį;
- atsargiai rinktis klaidų vertinimo metrikas esant nesubalansuotai duomenų imčiai;
- dauguma šioje dalyje pateiktų modelio prognozavimo tikslumo ir klaidų metrikų yra naudojamos ir tuo atveju, jei klasių skaičius daugiau nei dvi. Tačiau literatūros šaltiniuose yra pasiūlyta ir kitokių metrikų, taikytinų specialiai tokiems atvejams.