

Mašīnmācīšanās: Rītausmas agrā gaisma

Problēmas apraksts

Ivans saņēma neparastu pieprasījumu no potenciālā klienta: klasificēt 2D datu kopu trīs krāsās kodētās klasēs.

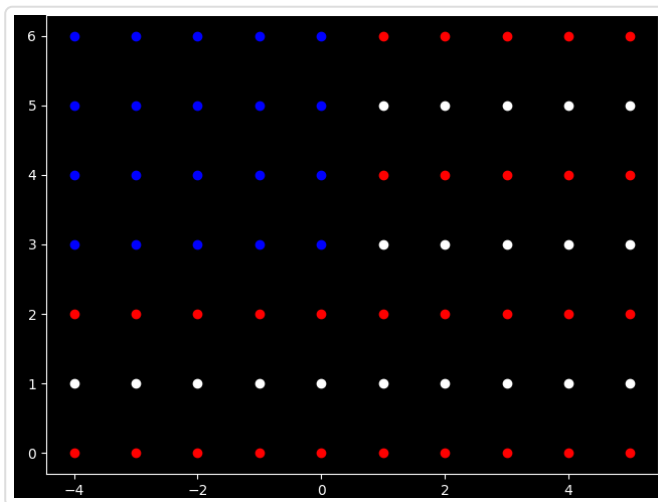
Kad viņš atvēra failu, Ivans lūkojās savā ekrānā. Datu punkti veidoja neparastu rakstu, kas izskatījās miglaini pazīstams...

Pēc brīža pārdomām viņš pasmaidīja. No visām klasifikācijas problēmām, ar kurām viņš bija saskāries, šī noteikti bija visradošākā.

Klientam bija specifiskas prasības: pirmkārt, noteikt, kuri klasiskās mašīnmācīšanās algoritmi teorētiski varētu sasniegt perfektu klasifikāciju bez jebkādam koordinātu transformācijām. Otrkārt, atrast minimālo matemātisko nosacījumu kopu, kas varētu nepārprotami atdalīt visas trīs klases.

Uzdevums

Jums ir dota 2D datu kopa, kā attēlots attēlā:



X ass ir horizontāla, Y ass ir vertikāla. Visi punkti atrodas veselu skaitļu koordinātēs. To krāsa apzīmē to klasi — tād ir balta klase, sarkana klase un zila klase.

Uzdevums {1}: Pieņemot, ka jebkādas koordinātu transformācijas ir aizliegtas, kuri no zemāk minētajiem klasifikatoriem teorētiski var iemācīties pilnībā atdalīt šīs trīs klases (prognozēt apmācības kopas (training set) apzīmējumus ar 100% precizitāti)?

- K Nearest Neighbours Classifier
- Support Vector Classifier
- Decision Tree Classifier
- K-Means Clustering
- Multi-Layer Perceptron

Uzdevums {2}: Šī uzdevuma mērķiem mēs definējam nosacījumu kā matemātisku apgalvojumu, kas sastāv tikai no koordinātām x un y , kā arī operatoriem $\{+, -, *, /, \% \text{ (modulo)}, <, \leq, =, >, \geq, \text{pakāpes operators}\}$, kas var tikt novērtēts kā patiess vai aplams. Piemēram, $2x - y/3 > 4$ vai $x^2 > -(y-x)^3$.

Izveidojiet šādu nosacījumu sarakstu tādā veidā, lai atbildot jā/nē uz katru nosacījumu, tas nepārprotami identificētu apzīmējumu, un implementējiet tos Python valodā.

Formāts

Pirmā daļa jāiesniedz kā jā/nē atbilžu saraksts, kas atbilst klasifikatoriem.

Otrai daļai iesniedziet Python failu, kas satur lambda funkciju sarakstu, kas reprezentē nosacījumus:

```
[lambda x, y: bool_expression_1, ..., lambda x, y: bool_expression_x]
```

Nepareiza, bet sintaktiski pareiza iesniegšana būtu:

```
[lambda x, y: 2*x - y / 3 > 4, ..., lambda x, y: x**2 > -(y-x)**3]
```

Katra no lambda funkcijām var iekļaut tikai `x, y, +, -, *, /, //, %, **, <, <=, ==, >=, >` un skaitliskas konstantes.

Svarīgi: Katrai funkcijai jābūt vienam salīdzinājumam (piemēram, `x + y > 5`). Jūs nevarat kombinēt vairākus salīdzinājumus ar aritmētiskajiem operatoriem (piemēram, `(a > b) * (c < d)` nav atļauts), jo tas radītu vairākas pareizi/nepareizi vērtības, nevis vienu.

Norādes

Ja Jūs nezināt, kā kāds vai visi šie klasifikatori darbojas, neuztraucieties. Jūs varat tos izpētīt tiešsaistē. Šeit ir noderīgs video, kas izskaidro katru no tiem (un vairāk), kas var būt labs sākumpunkts: <https://www.youtube.com/watch?v=E0Hmnixke2g>

Vērtēšana

- 25% par pareizu uzdevumu {1},
- 75% par pareizu uzdevumu {2}.

Uzdevums {1}

- pareizo atbilžu skaits / 5

Uzdevums {2}

- Par x nosacījumu skaitu, tiek piešķirts 1.776^{2-x} .
- Nepareizs formāts VAI nepareiza klasifikācija VAI python kļūda - 0