

Увод у вештачку интелигенцију

I колоквијум

24. 11. 2024

Задатак 1

Банка се суочава са проблем одласка клијената. Како би се зауставио губитак клијената банка жели да предвиди који ће клијенти напустити банку. Прикупљени су подаци о клијентима. Садрже информације као што су држава порекла, биланс рачуна, да ли поседује кредитну картицу итд. На основу датих података креирати моделе машинског учења којима се за једног клијента предвиђа да ли ће напустити банку или не.

Потребно је извршити следеће:

1. Учитати сет податка
2. Проверити да ли неке вредности података недостају. Уколико недостају, уклонити податке о клијентима за које постоје грешке у подацима.
3. Државе из које потичу клијенти кодирати методом One-Hot-Encoding. Ознаку пола записати методом Label- Encoding.
4. На основу датих података потребно је креирати моделе логистичке регресије користећи унакрсну валидацију података, $k = 5$.
5. У свакој итерацији унакрсне валидације оценити тренутни модел на тренинг и тестном скупу метрикама тачност и F1.
6. Одабрати најбољи модел на тестном скупу на основу метрике F1.
7. За одабрани модел, нацртати на истом графику ROC криве модела при предикцији на тренинг и тестном скупу. У легенди графика означити називе кривих.
8. Сачувати предикције модела са скупа за тестирање у csv формату. Излазни фајл треба да садржи податке у формату
 - а. ID клијента,
 - б. предикција моделом логистичке регресије,
 - с. стварна вредност.

Задатак 2

а) Имплементирати класу Просте линеарне регресије. Класа треба да садржи:

- методу за обуку модела на основу прослеђених сетова података која врши одређивање тежинских коефицијената w_0 и w_1
- методу за предикцију скупа података,
- методу за евалуацију скупа података и
- методу за преузимање тежинских коефицијената w_0 и w_1

Обука модела се врши методом најмањих квадрата. Имплементацију базирати на употреби NumPy библиотеке.

б) Проширити задатак под а) додавањем активационе функције на излаз модела. Предикција новог модела дата је према формули:

$$\varphi(x) = \text{AktivacionaFunkcija}(\text{LinearnaRegresija}(x))$$

Додатни материјал:

Додатни материјал који можете користити за решавање Задатка 2.

a) ReLU активациона функција:

$$ReLU(x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases}$$

b) Средња квадратна грешка:

$$MSE = \frac{1}{2m} \sum_{i=1}^m (y_i - \hat{y}_i)^2$$

c) Парцијални први извод функције грешке модела:

$$E'(w_1, w_0) = \begin{bmatrix} \frac{\partial E}{\partial w_1} \\ \frac{\partial E}{\partial w_0} \end{bmatrix} = \begin{bmatrix} \frac{1}{m} \sum_{i=1}^m ((w_1 x_i + w_0) - y_i) x_i \\ \frac{1}{m} \sum_{i=1}^m ((w_1 x_i + w_0) - y_i) \end{bmatrix}$$

НАПОМЕНА: Свој рад предати као *python* или *jupyter notebook* фајлове и сачувати на адреси
~\Desktop\Rad\Ime_Prezime_brojIndeksa\.