

# Домаћи задатак из предмета Основи машинског учења

## Цене кућа

25. новембар 2023.

### 1 Опис проблема

У оквиру овог домаћег задатка бавићемо се предвиђањем цена кућа на основу различитих атрибута. Ова врста задатка честа је у такмичењима на платформи [Kaggle](#). Циљ је развити модел машинског учења који ће, на основу доступних података, прецизно предвиђати цене некретнина.

Све потребне датотеке су доступне на [предвиђања цена куће](#). Доступан вам је скуп података у две датотеке: `train.csv` и `test.csv`. Прва датотека садржи податке о ценама кућа за обучавање модела (атрибути и стварне цене), док друга датотека садржи сличне податке, али без информација о ценама које ће бити предвиђане. Ваш задатак је направити модел који ће ефикасно предвиђати цене кућа на основу атрибута.

### 2 Опис скупа података

Атрибути у фајловима су следећи:

- **Id:** Јединствени идентификациони број некретнине.
- **MSSubClass:** Класификација куће према типу градње.
- **MSZoning:** Зонирање некретнине.
- **LotFrontage:** Дужина приступног пута некретнине.
- **LotArea:** Површина парцеле некретнине у квадратним метрима.
- **Street:** Тип пута ка некретнини (приступни пут).
- **Alley:** Тип алеје ка некретнини.
- **LotShape:** Облик парцеле некретнине.
- **LandContour:** Релјеф парцеле некретнине.
- **Utilities:** Доступни комунални сервиси.
- ... *(Остали атрибути)*
- **SalePrice:** Цена некретнине (само у `train.csv` фајлу).

Није нужно корисити све атрибуте из скупа података за обучавање модела, такође водити рачуна о томе да ли постоје сви подаци за сваку некретнину.

### 3 Предаја домаћег задатка

Након што обучите модел на основу `train.csv` података, потребно је генерисати предвиђања за некретнине из `test.csv` и сачувати их у `submission.csv`. Датотека `submission.csv` треба да садржи две колоне:

- **Id:** Јединствени идентификациони број некретнине из `test.csv`.
- **SalePrice:** Предвиђање вашег модела за цену некретнине.

Пример изгледа ове датотеке можете пронаћи у `sample_submission.csv`.

Потребно је предати `Submission.csv` као и комплетан код којим сте обучавали модел и генерисали предвиђања. Датотеке сачувати у оквиру директоријума под називом `Ime_Prezime_Broj_Indeksa`, а затим тај директоријум запаковати у `zip` датотеку са истим називом и предати на следећу адресу: `ml.fink2023@gmail.com`, са насловом “ОМУ-Предвиђања цена куће”.

**Напомена:** Датотеке предати искључиво у задатом формату.

Рок за предају је 25. децембар 2023. године у 23:59.