

A házi két részből áll. Az első részben meg kell oldani a notebook-ot. A második részben a notebook-ban elkészített függvényeket kell egy class-ba áttenni.

Első rész:

Minden cellában ki van kommentelve, hogy mit kell csinálni. A notebook-ban még nem kell osztályt készíteni, csak simán azokat a függvényeket elkészíteni amiket kér. Ha a komment nem kéri, hogy készíts függvényt akkor nem kell függvényt csinálni, hanem csak mondjuk plotolni, vagy megnézni az adat shape-jét. Azok a cellák amiben nem kell függvényt készíteni, csak neked lesznek segítségül abban, hogy jobban meg tudd érteni a modell működését/az adatokat.

Második rész:

A notebook-ban elkészített függvényeket szervezd össze egy osztállyá.

Osztály neve: KMeansOnDigits()

Konstruktor: Két paramétert várjon: n_clusters, random_state. Ezeket tedd példányszintűvé (self.n_clusters,)

Függvények:

load_dataset(): self.digits-be menti el a dataset

predict(): self.clusters-be menti el a predictált cluster osztályokat

get_labels(): self.labels-be menti el a labeleket

calc_accuracy(): self.accuracy-ba menti az acc-ot

confusion_matrix(): Itt készíts egy függvényt ennek, és csak az sklearn confusion_matrix()-nak a visszatérési értékét mentse el a self.mat-ba

Ha megnézik akkor semelyik függvény nem vár paramétereket, mindent az osztályon belül tárolunk el, és használunk fel.

A notebook-ot nem fogjuk értékelni csak az osztályt ami egy külön .py file-ban legyen, ahogy eddig is. Ha valaki nem akarja megcsinálni a notebook-ot, hanem egyből a class-t csinálja nem lesz semmi következmény, az csak azért van, hogy könnyebben tudjatok debug-ni.