

TÖL201G - Stærðfræði og reiknifræði

Verkefni 2

Fyrir þetta verkefni mega nemendur vinna í hópum, allt að þrír í hóp. Merkið lausnina með nöfnum allra í hópum og skilið einni hóplaun á gradscope. Þetta verkefni er umfangsmeira en fyrri vikublöð og því er nauðsynlegt að byrja snemma.

Skilafrestur er til **föstudagsins 21. apríl kl. 23:59**.

Skila þarf öllum kóða sem notaður er til að leysa verkefnið ásamt útskýringum. Gefið verður sérstaklega fyrir hversu auðvelt er að fylgja kóðanum. Athugið að minnstu fervika verkefni er hægt að leysa með `np.linalg.lstsq` fallinu í `numpy`.

MNIST

Í þessu dæmi munum við vinna með MNIST stafina aftur.

Gögnin eru 28×28 pixla myndir sem búið er að breyta í 784-vigra og er hver og einn merktur með réttu svari 0-9. Markmiðið er að læra reglu, þ.e. búa til flokkara sem tekur mynd sem inntak og gefur honum merkingu, þ.e. tölustaf 0-9. T.d. ef myndin er geymd í vigrinum x þá þarf að útfæra fall, f eða aðferð þar sem $f(x)$ skilar tölu á bilinu 0 til 9 og ætti helst að vera rétt.

Inntakinu er skipt í tvo hluta, A og B . Fyrri hlutinn A er notaður til að þjálf flokkarann f , og hér má nota bæði myndina og réttu merkinguna. Seinni hlutinn B er notaður til að meta hve góður flokkarinn f er.

A

Í notebook skemanu er búið að lesa inn gögnin og mynda $N \times 784$ fylki úr þeim (þ.e. hver mynd er línuvigur í fylkinu), þetta er gert til að einfaldara sé að vinna með aðferð minnstu fervika.

Þjálfðið 10 stykki tvíþátta flokkara $\tilde{f}_i, i = 0, \dots, 9$ með aðferðinni sem lýst er í kafla 14.2 í bók og mælið villutíðni fyrir hvern og einn. Myndið einn fjölþátta flokkara $\hat{f}$ með því að taka stærsta $\tilde{f}$ gildið sem niðurstöðu. Mælið villutíðni fyrir þann flokkara, bæði á þjálfunargögnunum og prófugögnunum. Í seinni liðum þarf bara að mæla villutíðni fyrir fjölþátta flokkarann á prufu og þjálfunargögnunum.

B

Endurtakið lið A með því að setja refsingu á hve stórt w er, bæði með því að nota (a) $\|w\|^2$ og (b) $\|D_h w\|^2 + \|D_v w\|^2$. Veljið λ á bilinu $10^{t/2}$ fyrir $t = -4, \dots, 4$ og athugið villutíðni. (Sjá kafla 15.3.3 fyrir notkun og skilgreiningu á D_h og D_v).

C

Bætið við nýjum breytum með því að velja vigra r_j sem handahófskennda vigra með gildin ± 1 og tölur s_j , $\max(r_j^T x + s_j, 0)$ (sjá kafla 13.2 og 14.2.2 í bók). Fyrir hverja mynd bætið þið við k slíkum slembigögnunum, þjálfðið flokkara eins og í lið A og mælið villutíðni. Keyrið þetta fyrir $k = 100, 500, 1000, 10000$.

D

Bætið við nýjum breytum með því að mæla hækkun milli aðlægra pixla, ef i, j eru tvö hnit sem eru aðlæg (annað hvort til hliðar eða upp og niður) þá búum við til nýja breytu $z_{i,j} = \max(x_i - x_j, 0)$. Hversu margar breytur verða í þessu módeli og hver er villutíðnin?

E

Notið breytur úr liðum C (besta gildi á k að ykkar mati) og D (saman) ásamt því að velja nota $\|w\|^2$ refsingu fyrir aukabreyturnar (en ekki fyrstu 784 breytur sem eru beint úr myndinni). Mælið villutíðni