

Introduction to machine learning

Introduction to machine learning, Extra project

Ensimmäisenä tehtävä oli luoda perceptron algoritmi, jota sitten käytetään käsinkirjoitettujen numeroiden luokitteluun. Tässä tapauksessa tarkoituksena oli luokitella ainoastaan 1 ja 0. Tein toteutuksen Matlabilla. Tein numeroiden lataamiseen oman scriptin, joka valitsee numeroista vain 1 ja 0, ja jakaa pikselidatan sekä oikeat luokat omiin muuttujiinsa. Tämän lisäksi on itse perceptron funktio omassa tiedostossaan, joka hoitaa luokittelijan opettamisen. Sitten on vielä classifier funktio, jolla voi luokitella käyttäen aiemmin luotua luokittelijaa.

Ensimmäisenä tein 2-ulotteisia testiaineistoja, jotta tiedän toimiiko luokittelija. Kaikissa seteissä pisteitä oli sen verran vähän että jos luokat olivat eroteltavissa, niin algoritmi konvergoi parissa iteatiossa. Seuraavalla sivulla muutamia kuvia, joista näkee minkälaisia ryhmiä luokittelija sai luotua. Kohtasin pieniä ongelmia siinä miten olisin saanut MatLabissa esitettyä pisteet omina luokkiina, mutta tässä tapauksessa onneksi pisteitä on niin vähän, että on melko selvää mitkä pisteet kuuluvat mihinkin luokkaan. Tarkastin kuitenkin käsin, että saadut tulokset olivat järkeviä.

Seuraavaksi latasin apufunktiolla 5000 datapistettä mnist aineistosta ja otin sieltä vain 1 ja 0. Opetusaineistoon ja testiaineistoon tuli kumpaankin noin 500 alkia. Perceptron algoritmi konvergoi toisella iteraatiolla. Tämän jälkeen kun luokittelijan antoi syötteenä testiaineistolle, oli virheprosentti 0%. Tulos ei sinällään ollut kovin yllättävä, koska aineiston dimensio oli todella suuri, sekä 1 ja 0 ovat melko kaukana toisistaan jos asiaa miettii pikselitasolla.

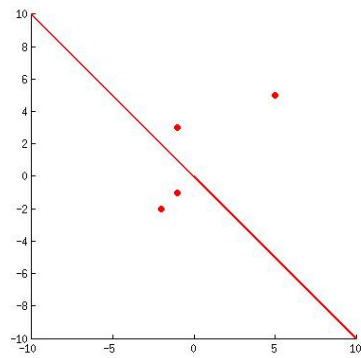


Figure 1: Luokittelijan testi 1

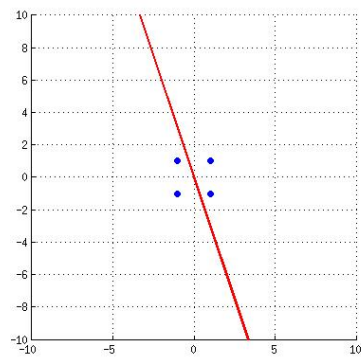


Figure 2: Luokittelijan testi 2

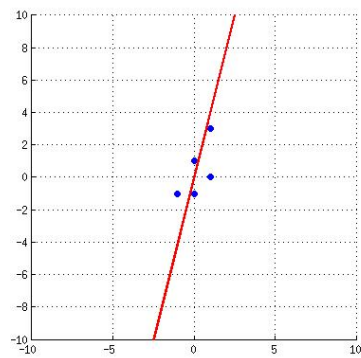
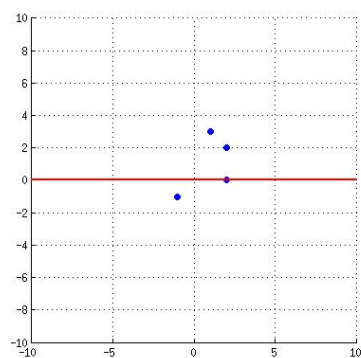


Figure 3: Luokittelijan testi 3, Kaikista toimivista luokista tässä oli pienin marginaali.



2

Figure 4: Pisteet eivät olleet luokiteltavissa, algoritmi ei konvergoanut