

*Nämä ovat viimeiset viikkoharjoitukset. Vapun aattona 30.4. ei ole luentoja. Kurssin tentti on maanantaina 14.5. klo 9–12.*

1. Tarkastellaan hypoteesikokoelman oppimista käyttäen enemmistöäänestystä opittujen  $M$ :n hypoteesin kesken. Olkoon kaikkien hypoteesien virhe  $\epsilon$  ja hypoteesien tekemät virheet riippumattomia toisistaan. Laske oppimisalgoritmin tekemä virhe parametrein  $M$  ja  $\epsilon$  ilmaistuna. Laske virhe kun  $M = 5, 10, 20$  ja  $\epsilon = 0.1, 0.2, 0.4$ .
2. Mikä on 6-bittisten Boolean käsitteiden otosvaativuus? Paljonko esimerkkejä tarvitaan, kun opitun hypoteesin sallitaan tekevän virhettä vain 1% todennäköisyydellä 0.999? Paljonko erilaisia kohdekäsitteen esimerkkejä on olemassa?
3. Osoita, että  $k$ -päättöslistoin voidaan esittää kaikki (Boolean) funktiot, jotka voidaan esittää korkeintaan syvyyttä  $k$  olevilla päätöspuilla.
4. Tarkastellaan *monotonisten monomi*en l. muuttujakonjunktioiden polynomista oppimista. Olkoot muuttujat  $x_1, \dots, x_m$ . Seuraava menetelmä on PAC-oppimisalgoritmi tälle hypoteesiluokalle.

(a) Aseta hypoteesiksi  $h = x_1 \wedge x_2 \wedge \dots \wedge x_m$

(b) Positiivisella esimerkillä  $(a, 1)$ , jos  $a_i = 0$  poista  $x_i$  hypoteesista  $h$

Joka vaiheessa  $h$  sisältää kaikki kohdekäsitteen  $\gamma$  muuttujat (+ muita). Näin ollen, (lopullinen) hypoteesi ei tee virhettä negatiivisen esimerkin luokittelussa.

Analysoi menetelmän otosvaativuutta ja osoita siten, että sen aikavaativuus on polynominen tarvittavien parametrien suhteen.

*Ohje:* Tarkastele muuttujaa  $z$ , joka esiintyy  $h$ :ssa, muttei  $\gamma$ :ssa. Jokainen  $h$ :n tekemä virhe voidaan laskuttaa ainakin yhdelle tällaiselle muuttujalle  $z$ .

Muuttujan  $z$  sanotaan olevan haitallinen, jos sen virhe on vähintään  $\epsilon/m$ . Mikä on  $h$ :n virhe, jos se ei sisällä haitallista muuttujaa? Mikä on todennäköisyys sille, ettei haitallista muuttujaa  $z$  ole poistettu  $h$ :sta, kun on nähty  $n$  esimerkkiä? Vaadi todennäköisyyden olevan pienempi kuin  $\delta$ .