



Haaga-Helia

Lili Aunimo, FT, yliopettaja

Haaga-Helia ammattikorkeakoulu

Pilvityökalut ja tekoäly datan
yhdistämisessä ja
valmistumisaikaan
vaikuttavien tekijöiden
analysoinnissa

Johdanto

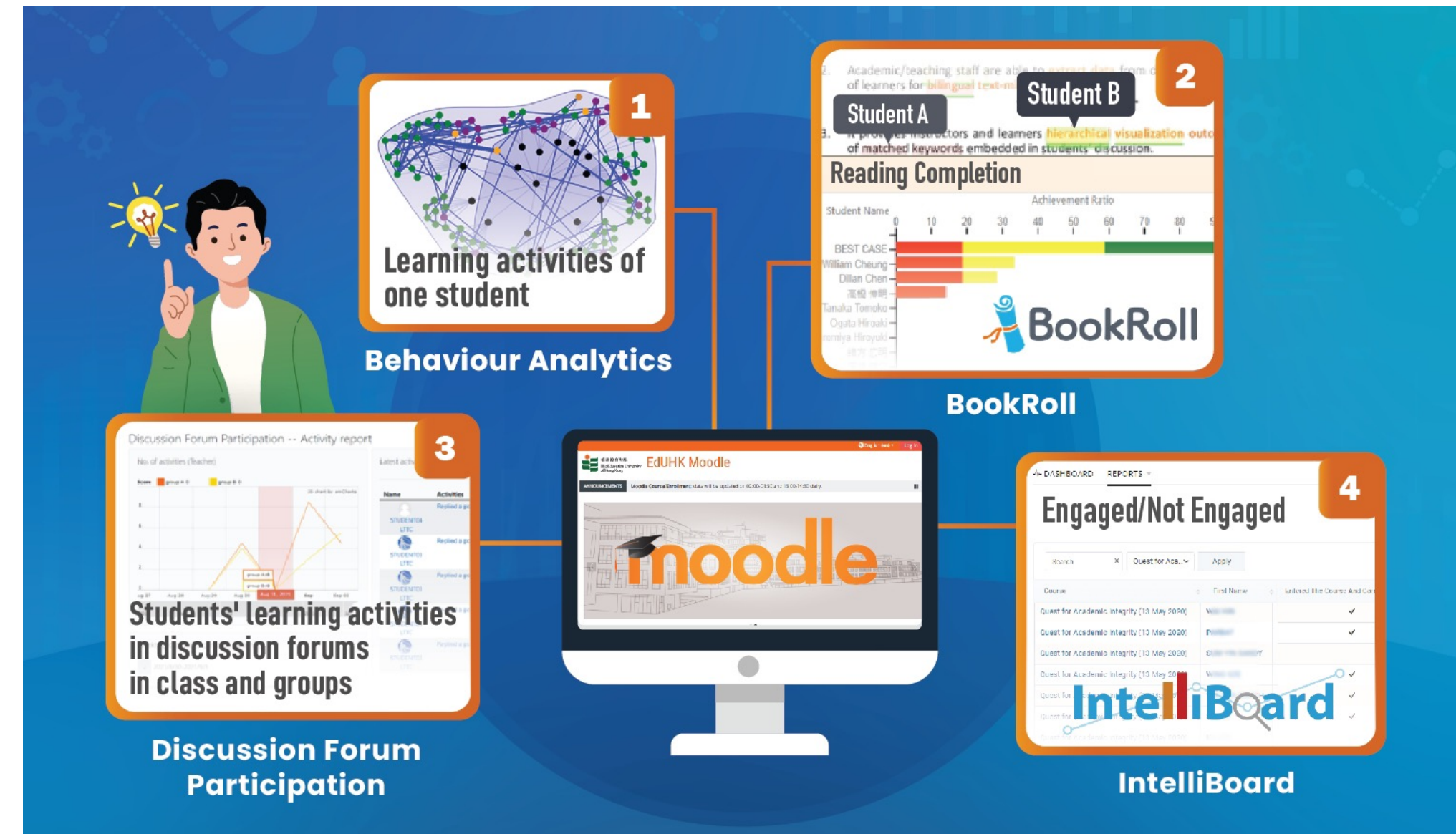
Digitalisaation myötä opiskelijan oppimisprosessista on paljon dataa.

Datasta voidaan louhia:

1. opintopolkuja,
2. opiskelutapoja
3. pullonkauloja,
4. jne.

Rakensimme prediktivisen mallin opiskeluajan ennakointiin.

Selvitimme, mitkä piirteet ennakoivat nopeaa valmistumista.



Source: <https://www.lttc.eduhk.hk/moodle/learning-analytics/>

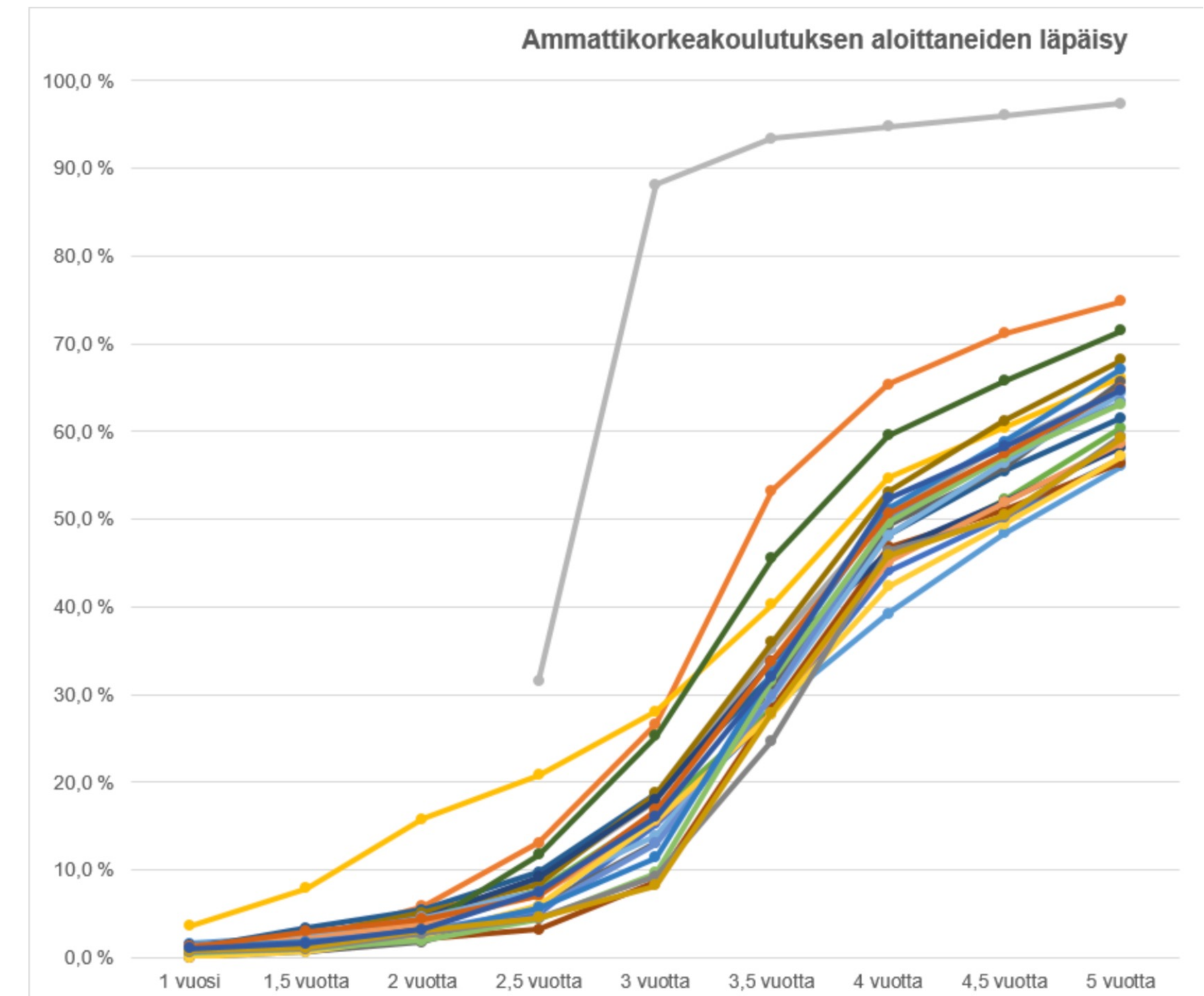
Tutkimuskysymykset

1. Voidaanko oppilaitoksessa rakentaa oppimisanalytiikkaa hyödyntävä prediktiivinen malli? Jos kyllä, miten?
2. Millaiset **oppimisympäristöt, -prosessit ja työvälineet** edesauttavat oppimista?
3. Millaiset **sisällöt ja pedagogiset käytännöt** parantavat oppimistuloksia?

Parempi oppiminen voi myös nopeuttaa opintoja.

Paremmasta ja tehokkaammasta oppimistesta hyötyy opiskelija, korkeakoulu ja yhteiskunta.

Graafi



Vipunen.fi Ammattikorkeakoulutuksen aloittaneiden läpäisy. Aloituslukuvuosi: 2015/2016.

Aikaisemmat tutkimukset

Keskittyvät datan analysointiin vain yhdestä kurssista

MOOC:it ja verkkokurssit ovat yleisimpiä tutkimuskohteita

Tässä tutkimuksessa analysoitiin:

1. dataa sekä Pepistä että Moodlesta , ja
2. koko opintopolun ajalta.

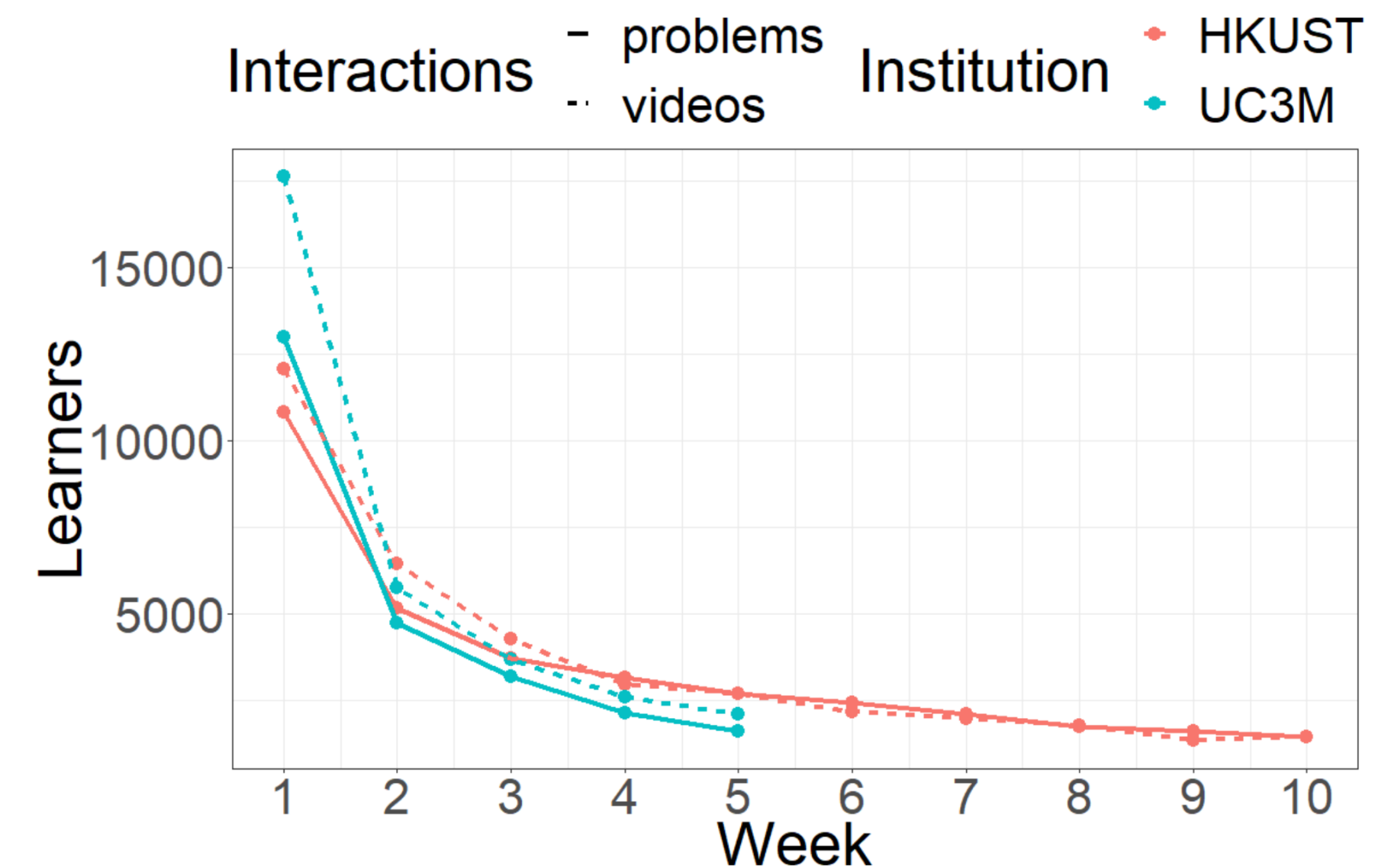


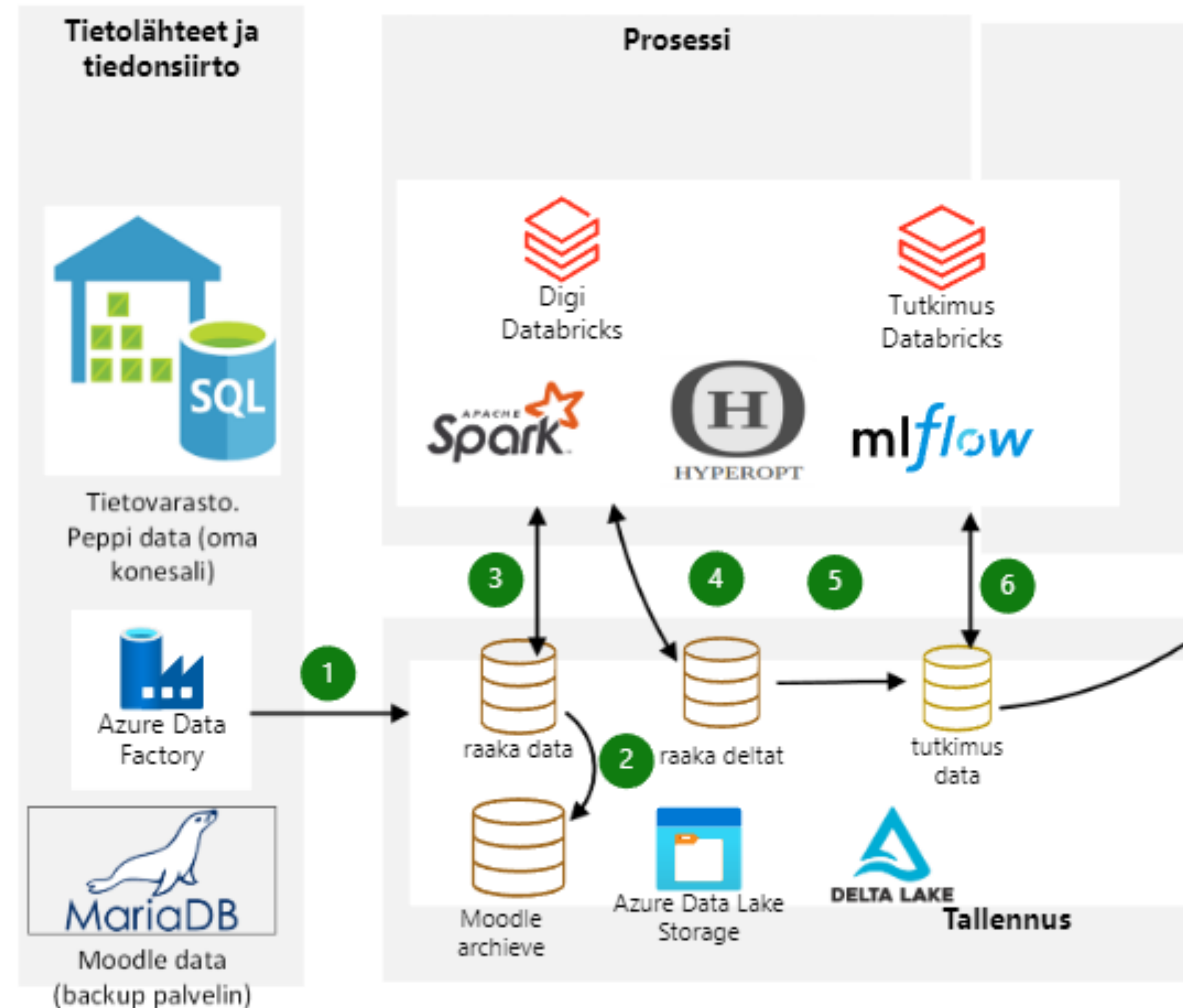
FIGURE 1. Evolution of the number of learners engaging with videos and exercises.

Moreno-Marcos et al. (2020). Analysis of the factors influencing learners' performance prediction with learning analytics. IEEE Access, 8, 5264-5282.

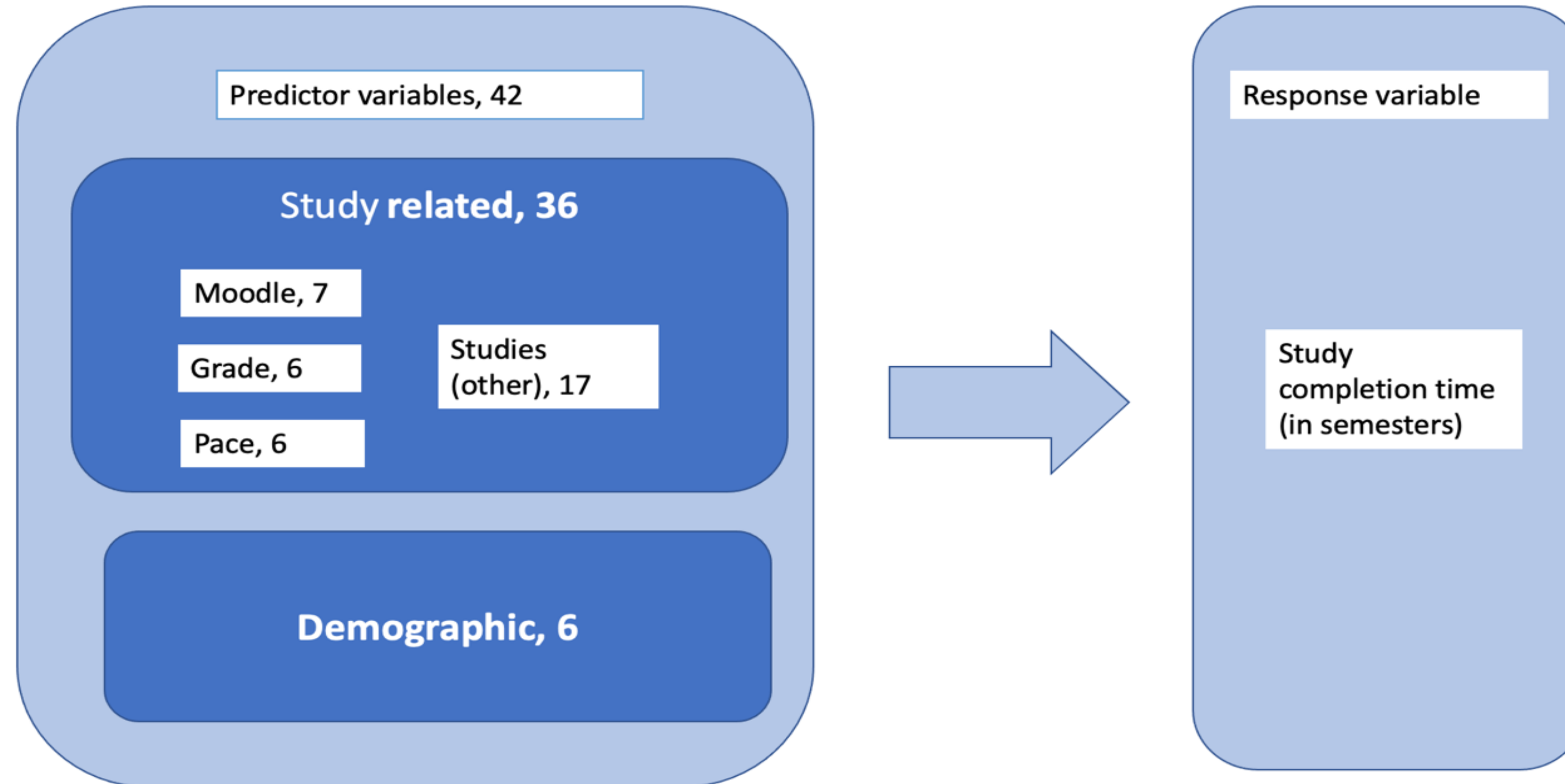
Data

Datan prosessikaavio (Peppi, Moodle, Azure pilvi)

- Datalähteinä Peppi ja Moodle
 - HH:sta valmistuneet opiskelijat vuosina 2014-2021 (alempi amk)
 - 12683 opiskelijaa
 - 82 muuttujaa (piirrettä), joista 43 mukana mallissa
- Datan koostaminen ja esikäsittely työläs prosessi
 - Muuttujissa tulkinnanvaraisuutta ja vaihtelua vuosien varrella
 - Moodle datassa paljon puuttuvia arvoja



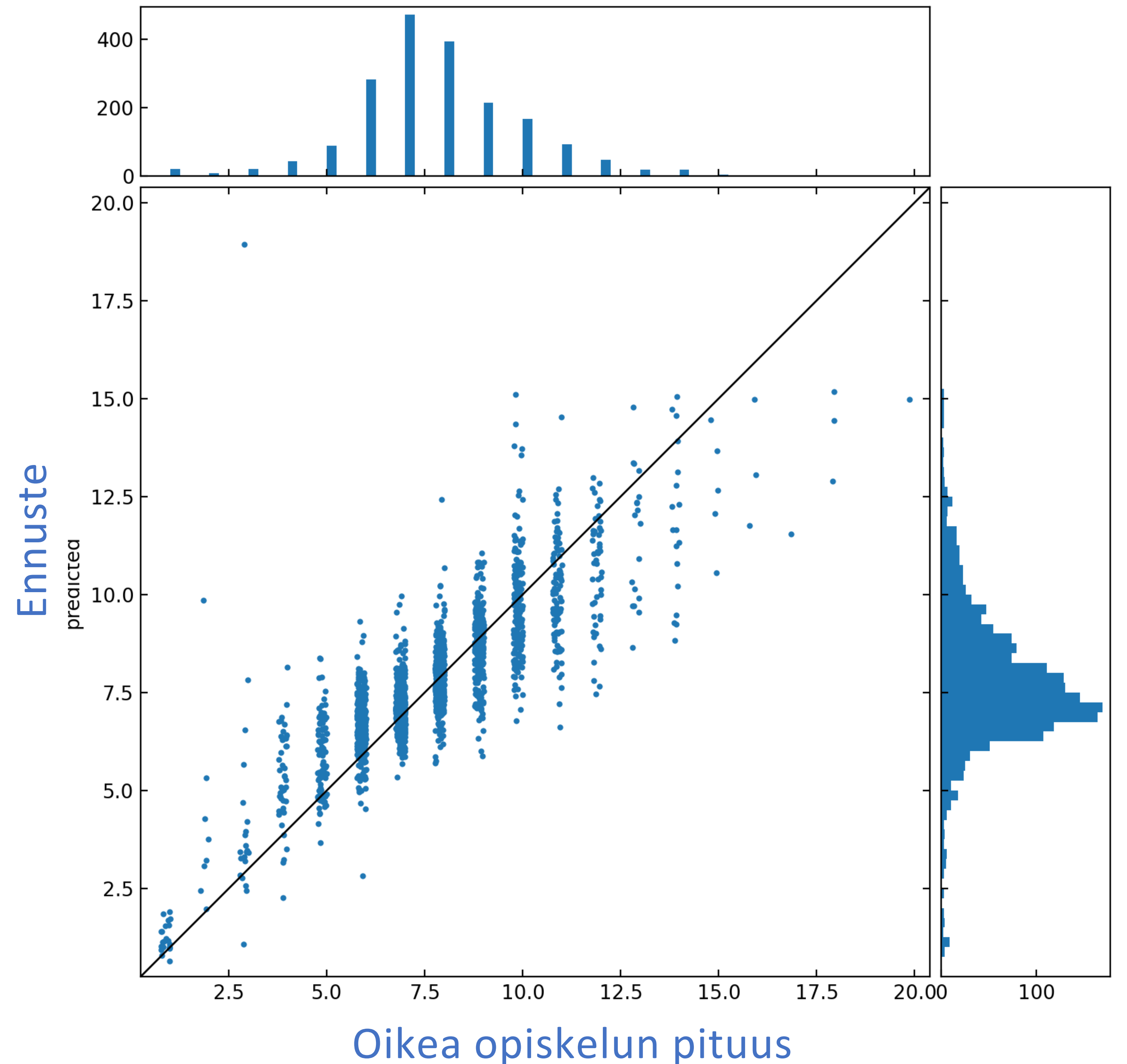
Mallin muuttujat



Malliin valittiin 43 muuttujaa. Ennustettava muuttuja oli opintojen kesto. Muut muuttujat voidaan jakaa 5 luokkaan: Moodle, arvosana, tahti, muut opintointoihin liittyvät ja demografiset.

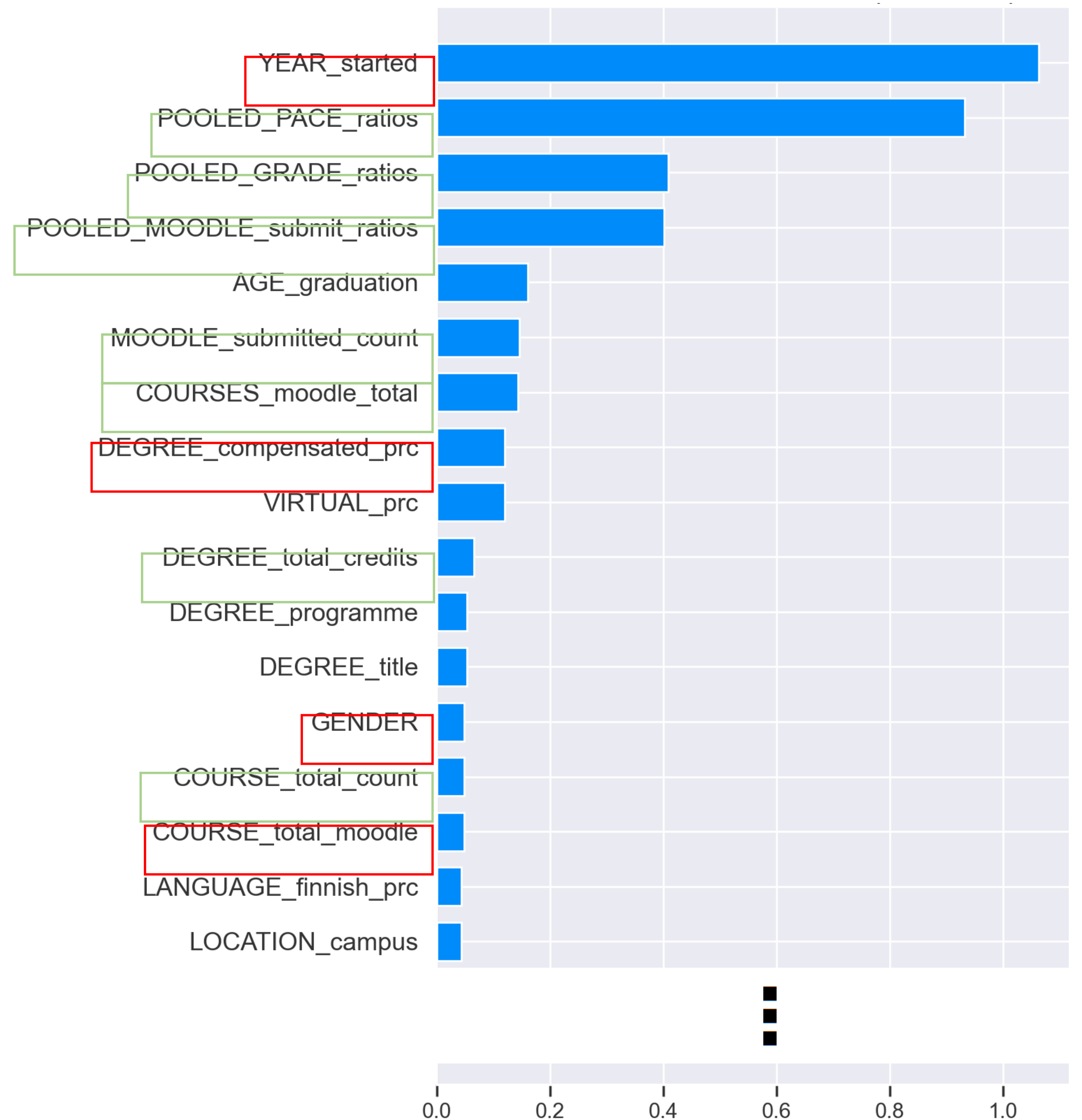
Datan mallinnus

- Regressioanalyysi CatBoost gradient boosting –mallilla
- Datasta 85% opetukseen, 15% testaukseen
- Lopullisen mallin tarkkuus testidatalle jopa $0.70 R^2$
 - Selittää $\sim 70\%$ vaihtelusta



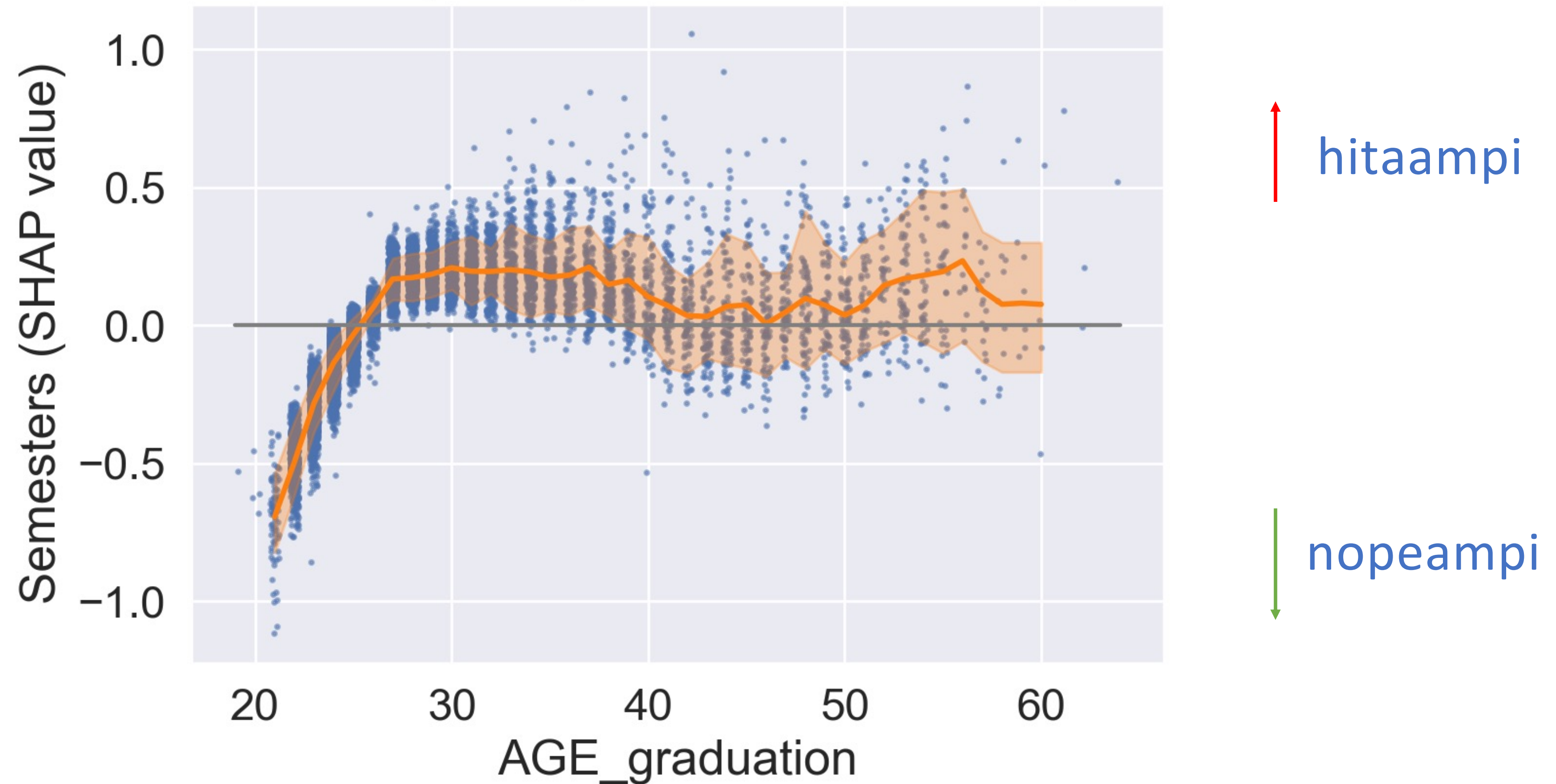
Tärkeimmät piirteet

- Piirteiden painoarvoa mitattiin SHAP menetelmällä*
- Peliteoriaan perustuva matemaattinen arvio yksittäisen piirteen vaikutukselle lopputulokseen
 - piirre ~ ”pelaaja”
 - mallinnus/näyte ~ ”peli”



Opiskelijan iän vaikutus

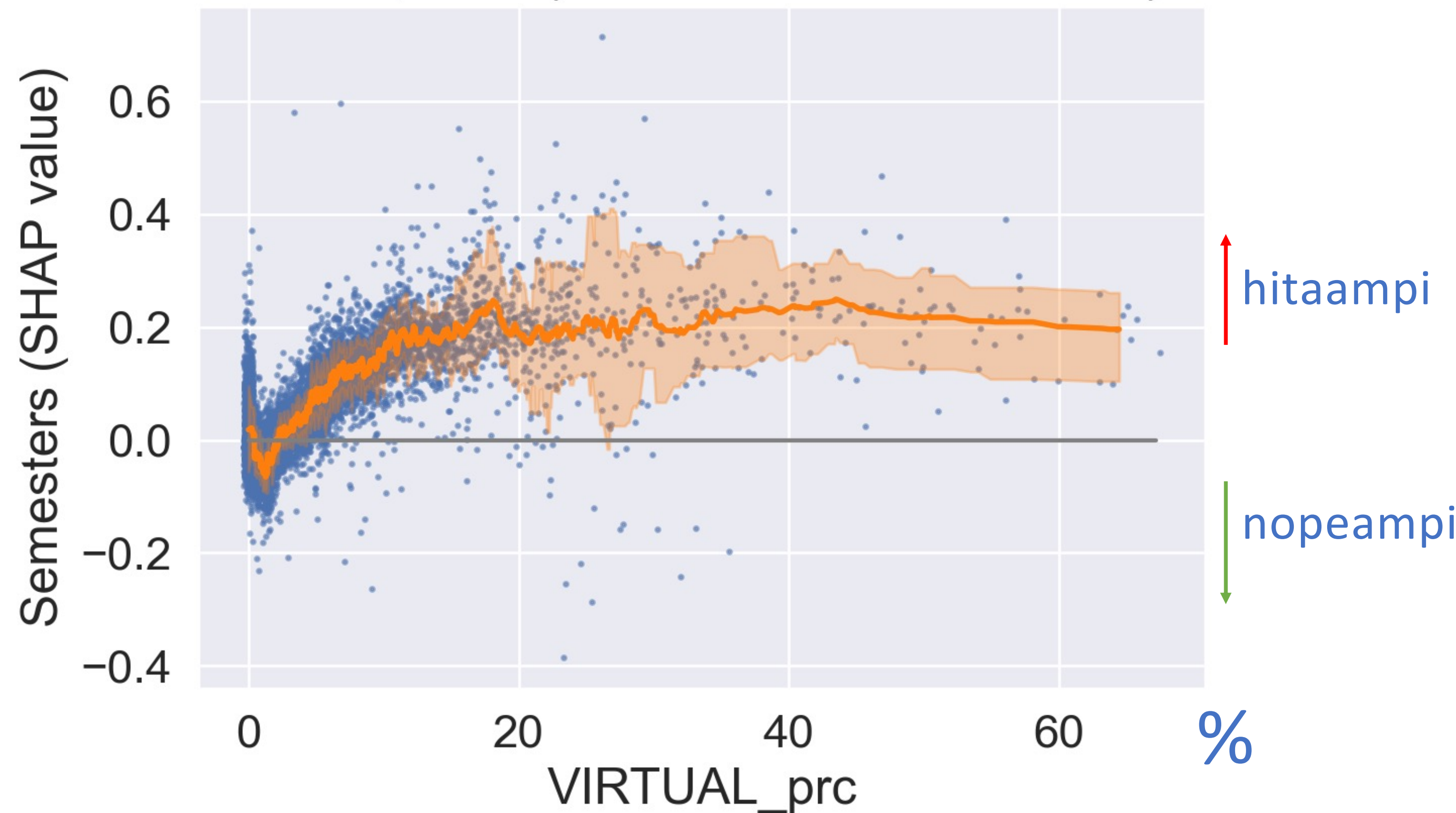
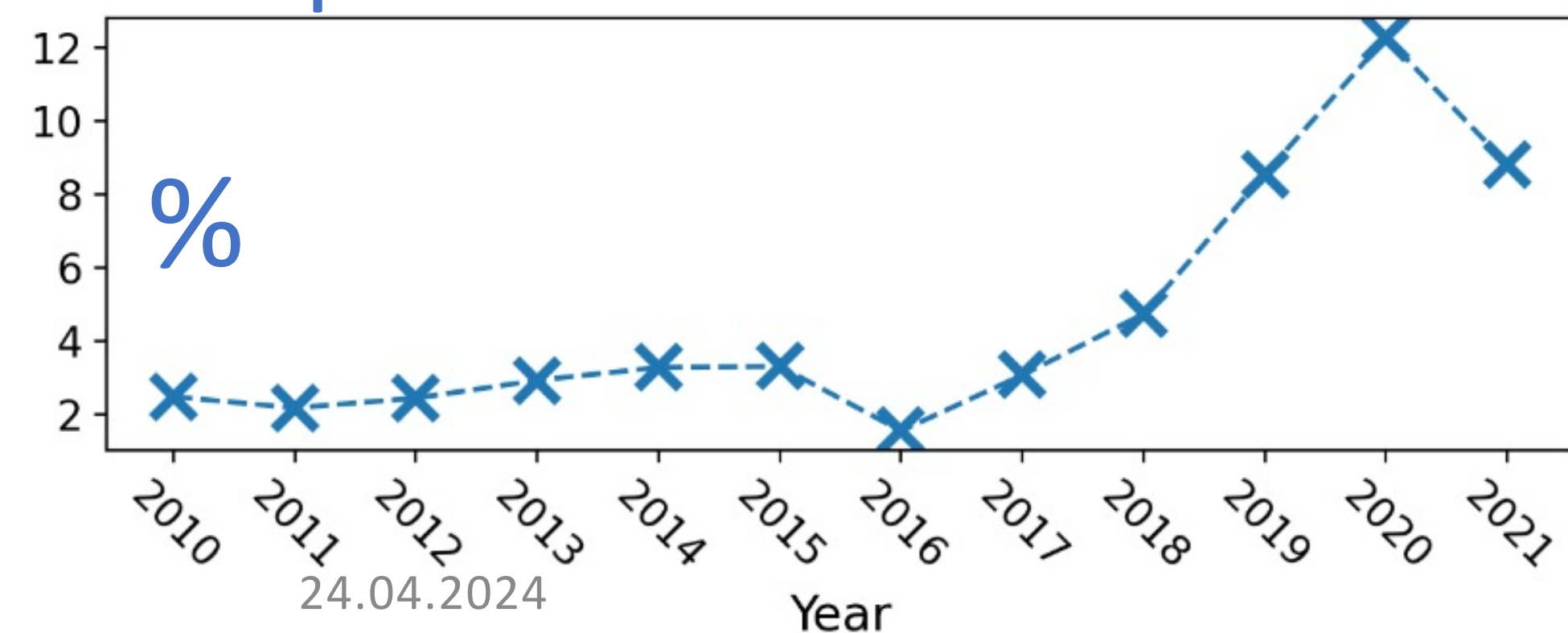
- Mitä nuorempi, sitä nopeampi valmistuminen
- 30 - 40v hitain ikäluokka



Etäopetuksen vaikutus

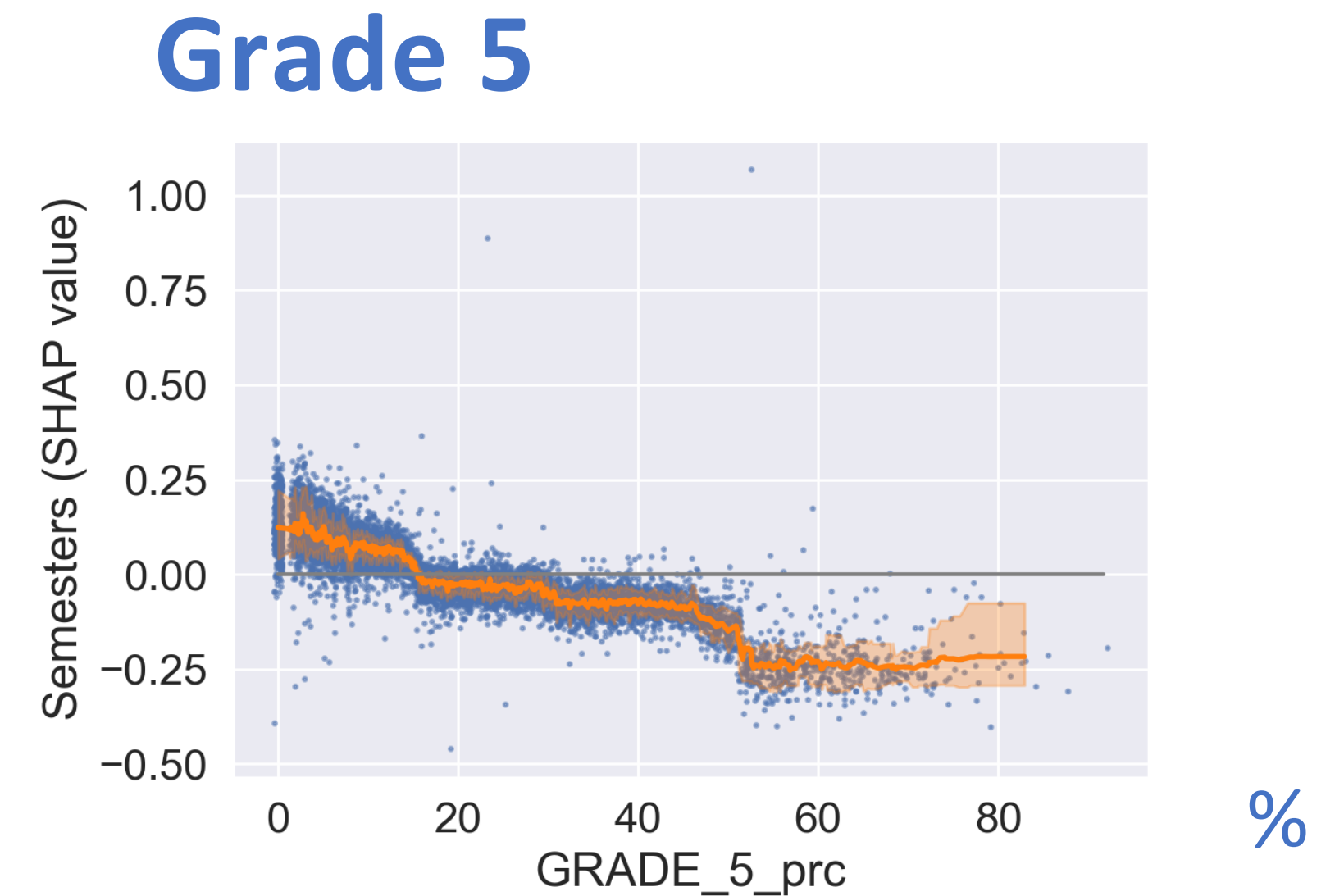
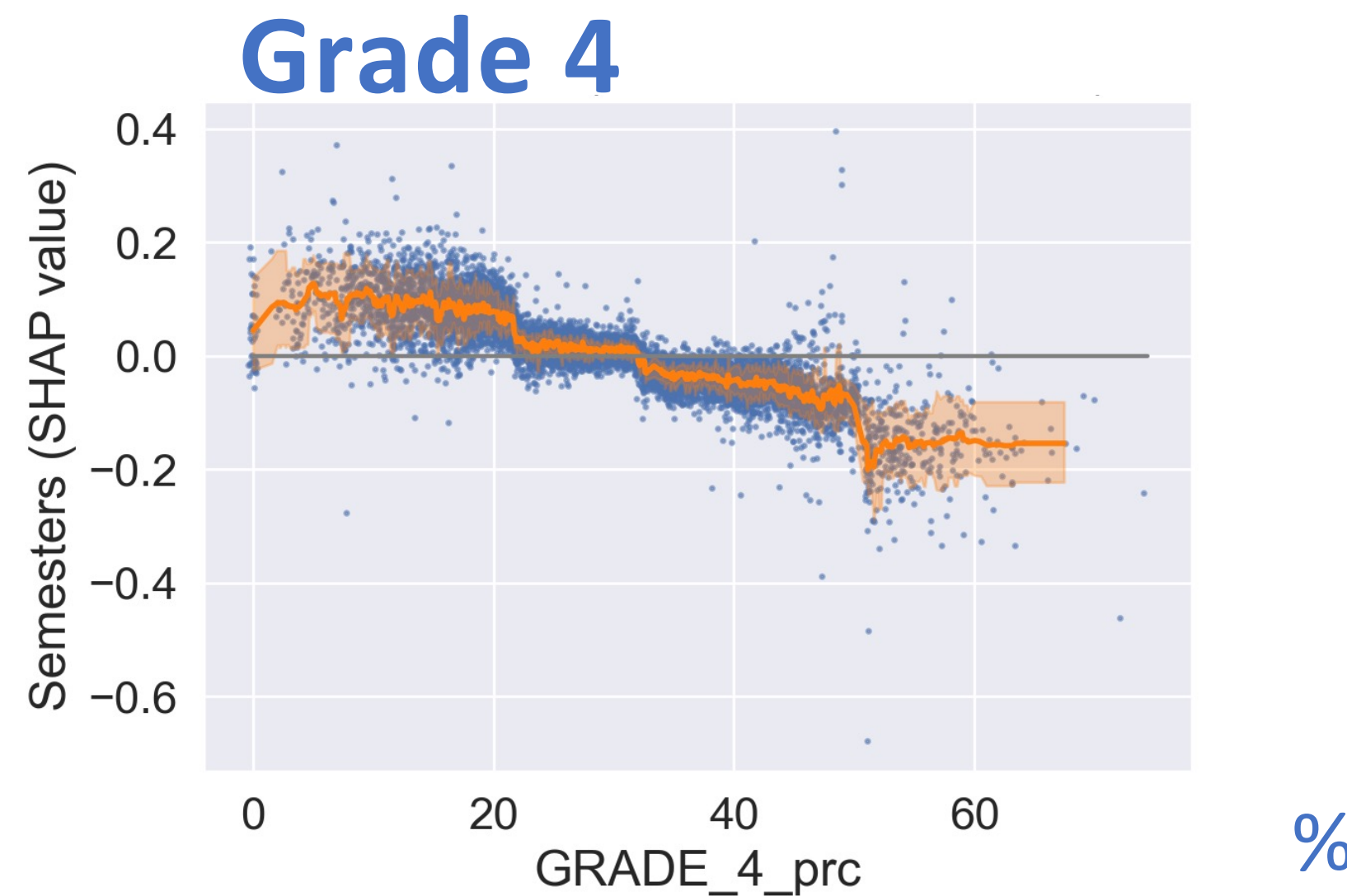
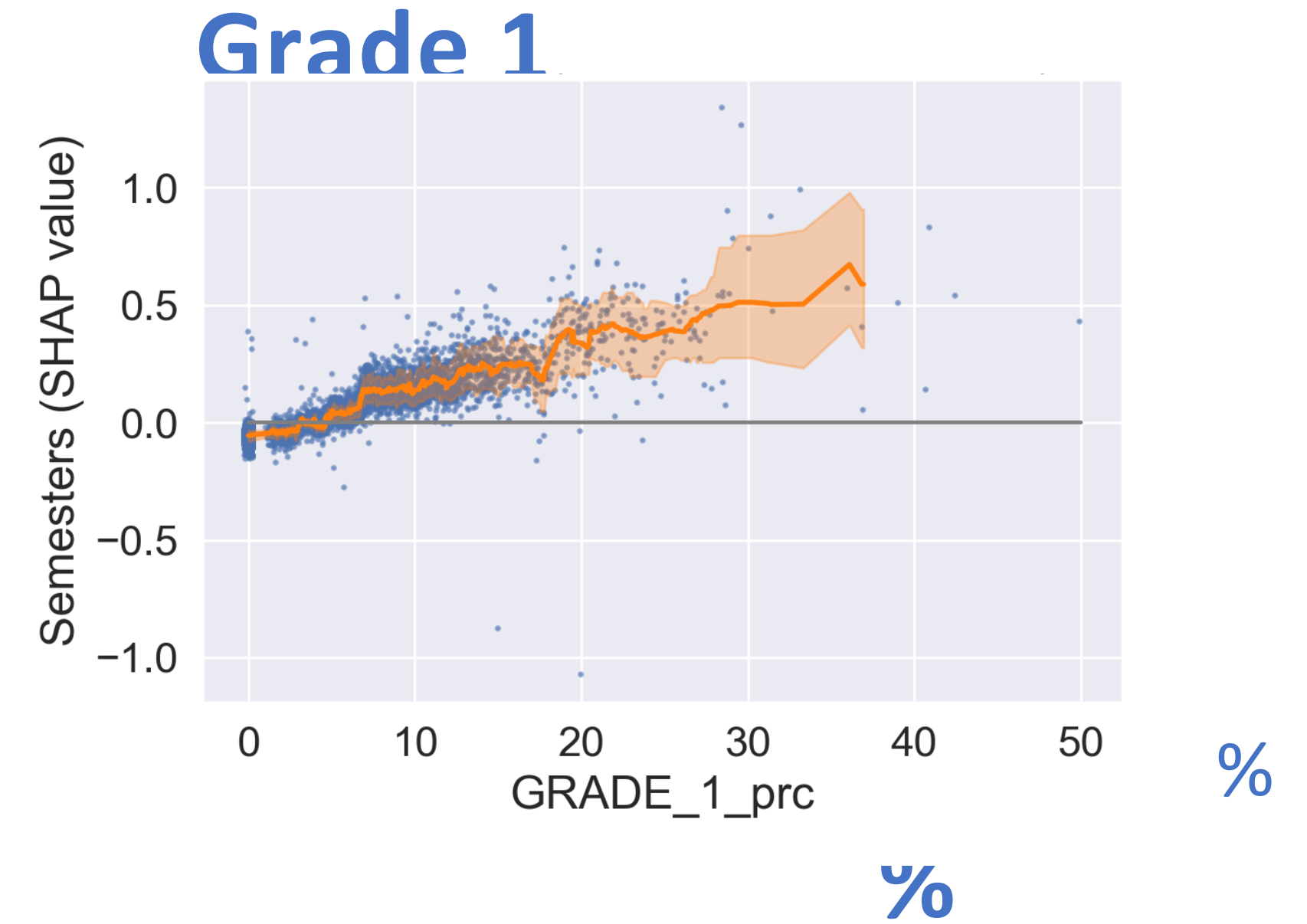
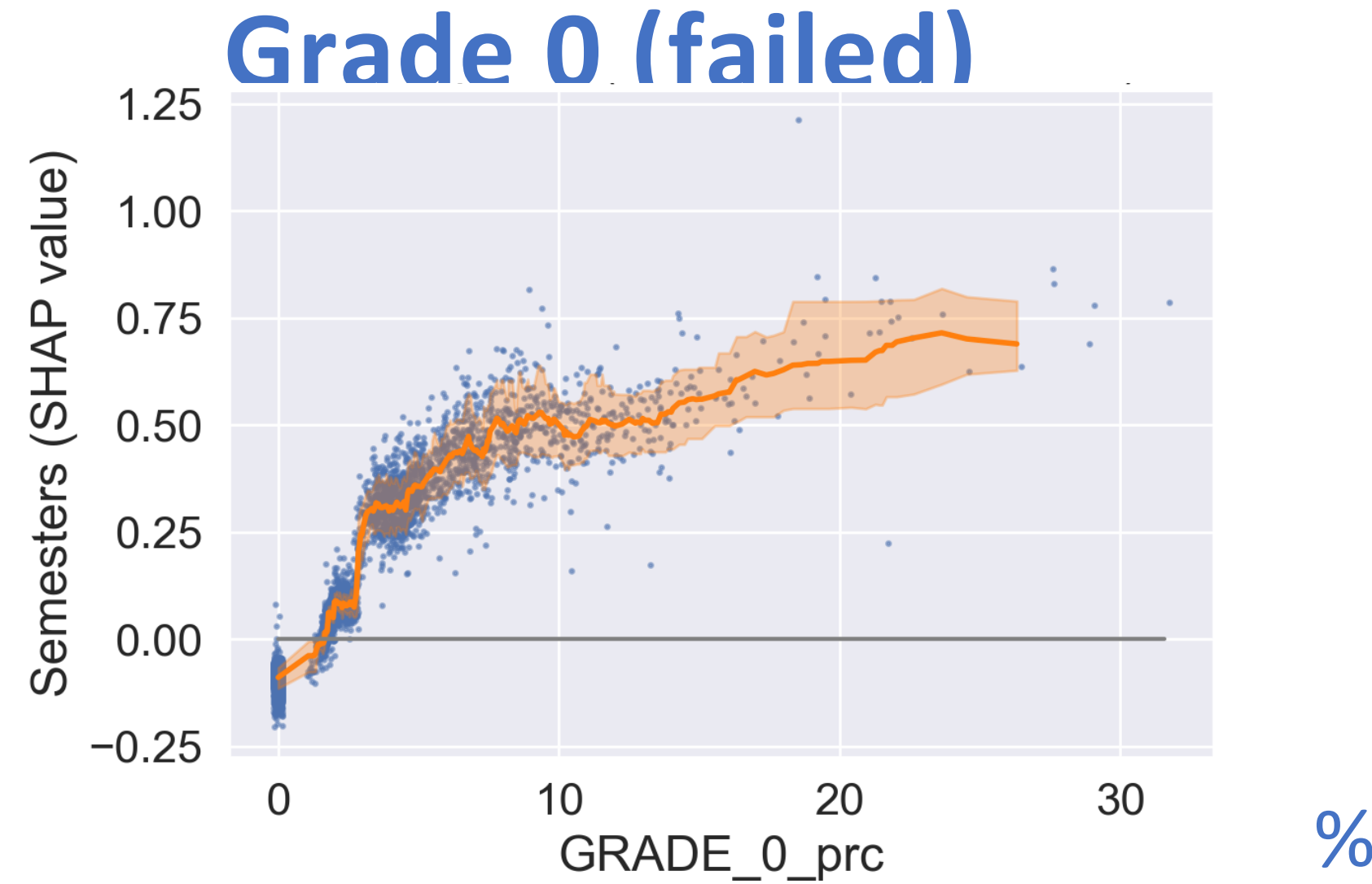
- Prosenttiosuus opinnoista
- Yli >3%
etäopetusta hidastaa valmistumista
- Huom: COVID-19:n kokonaisvaltaiset vaikutukset ei huomioitu

Etäopiskelu HH:ssa



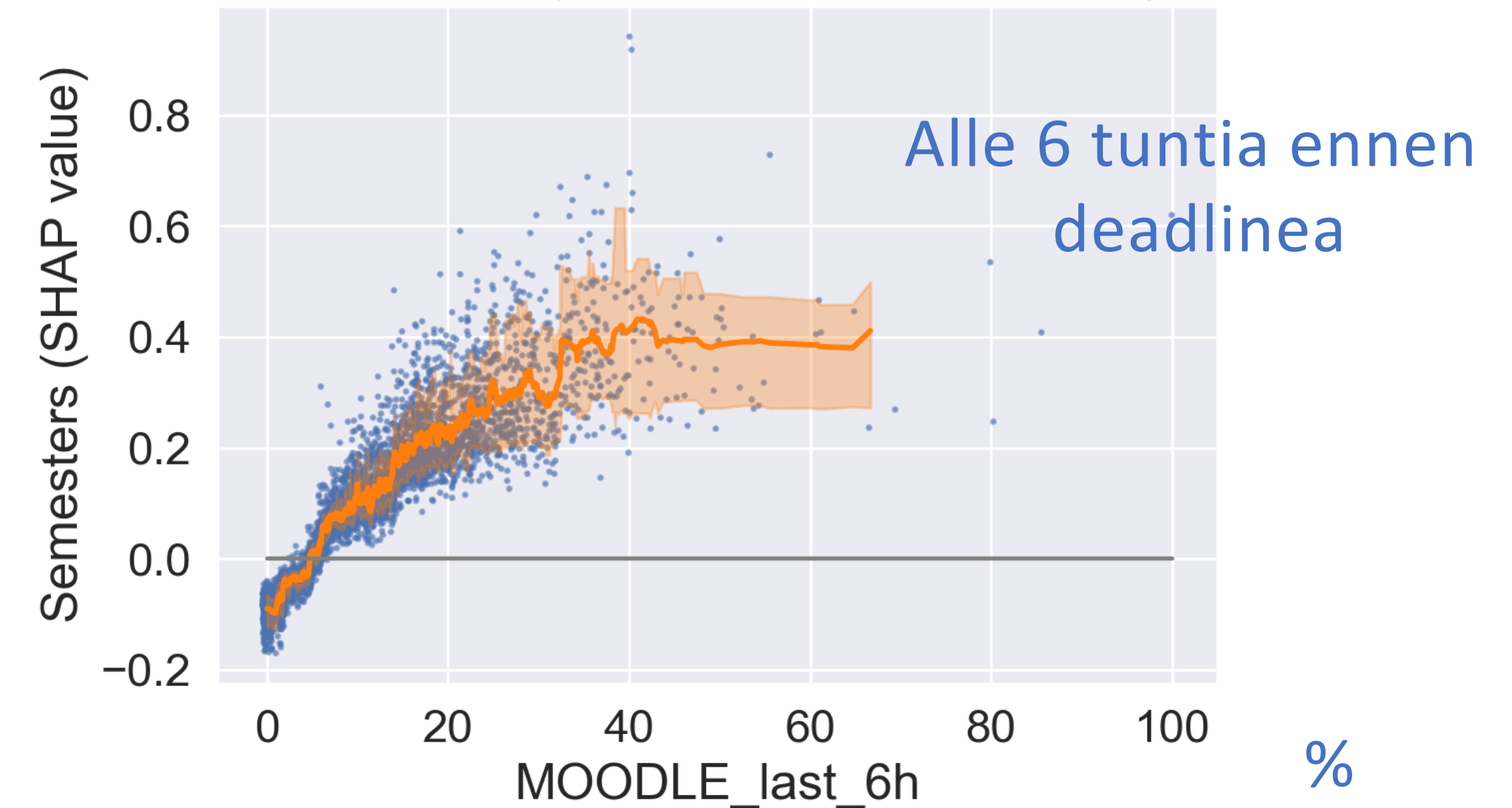
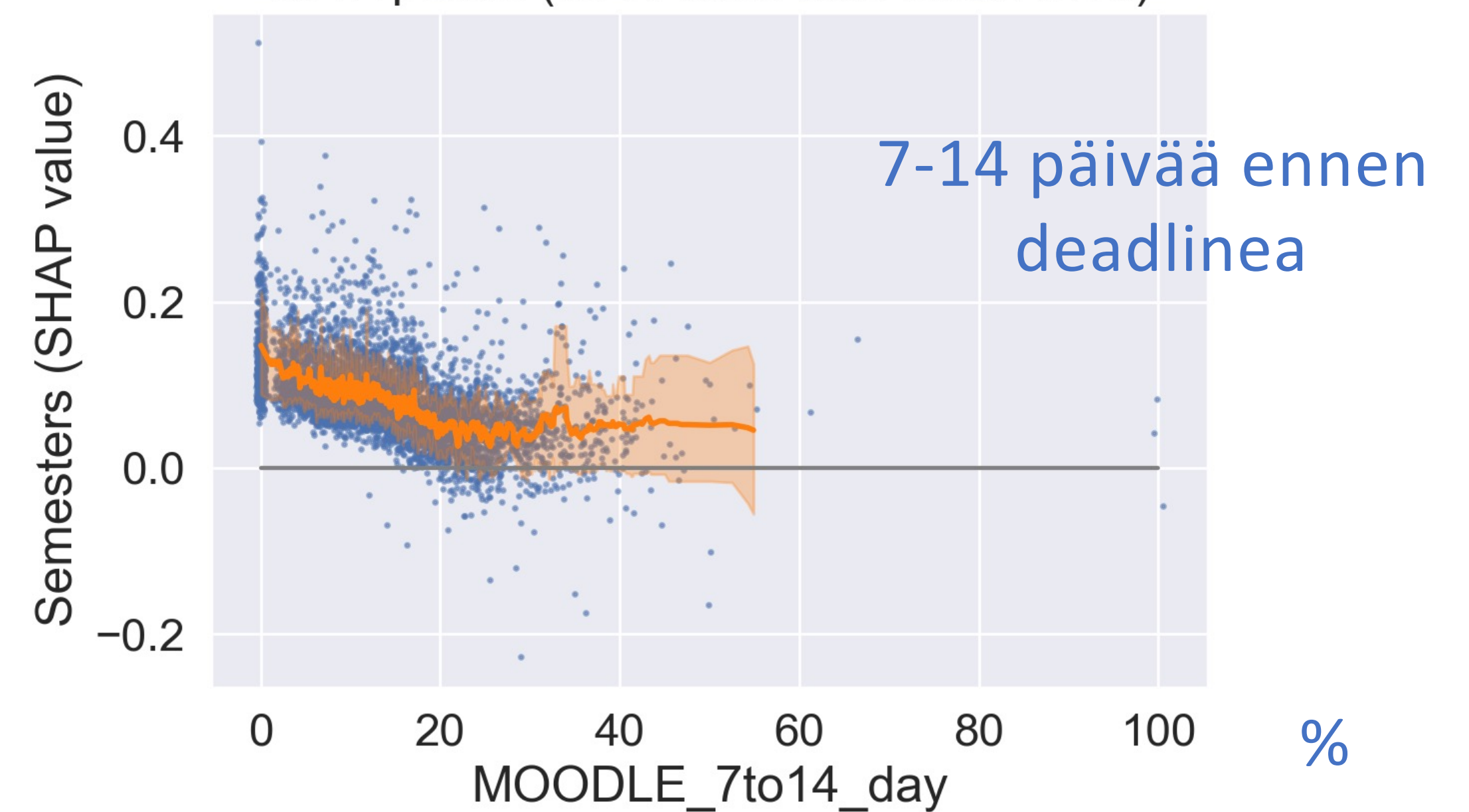
Arvosanojen vaikutus

- Prosenttiosuus arvosanoista
- Paremmat arvosanat, nopeampi valmistuminen
 - Mutta ei liian hyvät!
- Jo pieni määrä (>5%) heikkoja arvosanoja vaikuttaa opintoja pidentävästi



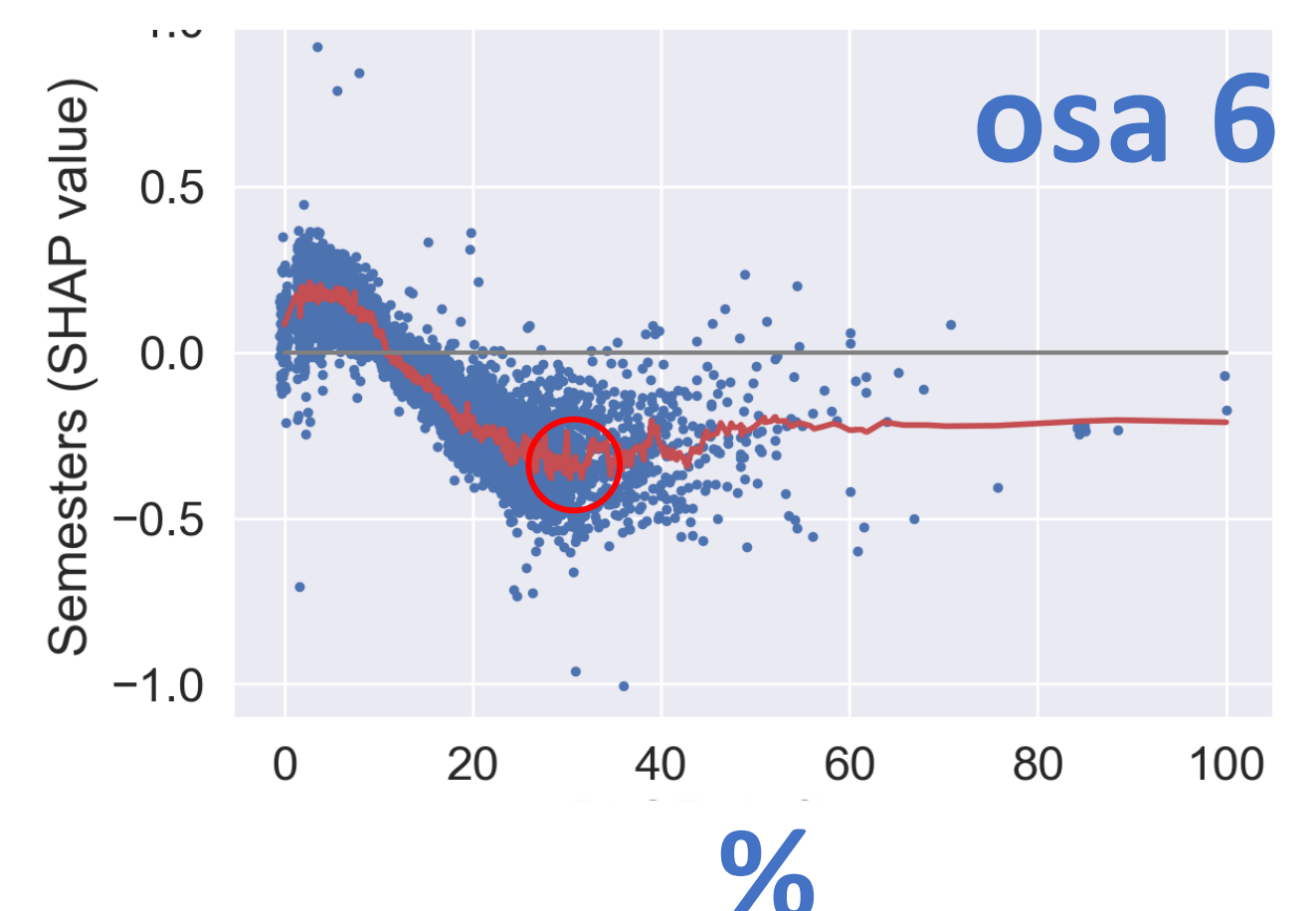
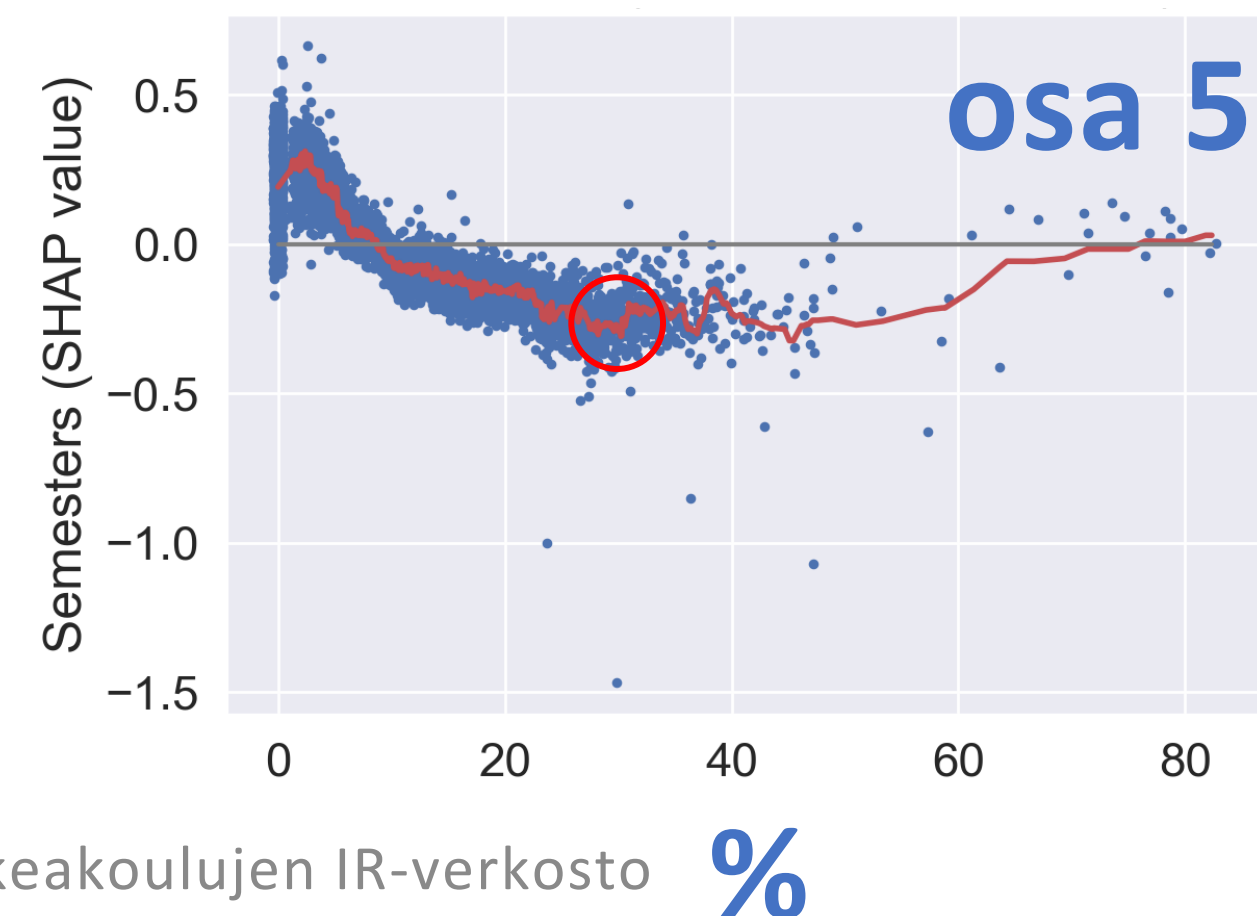
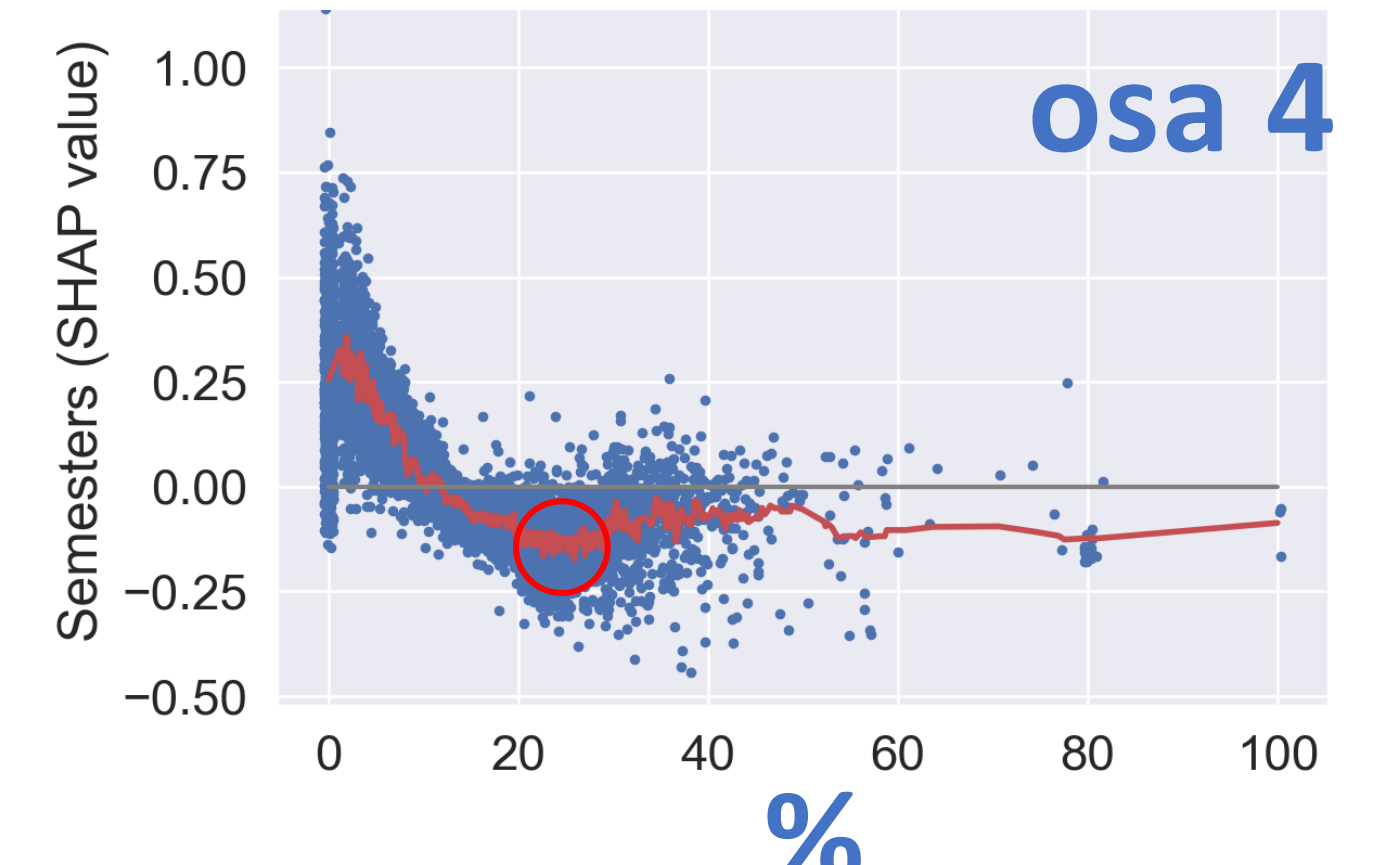
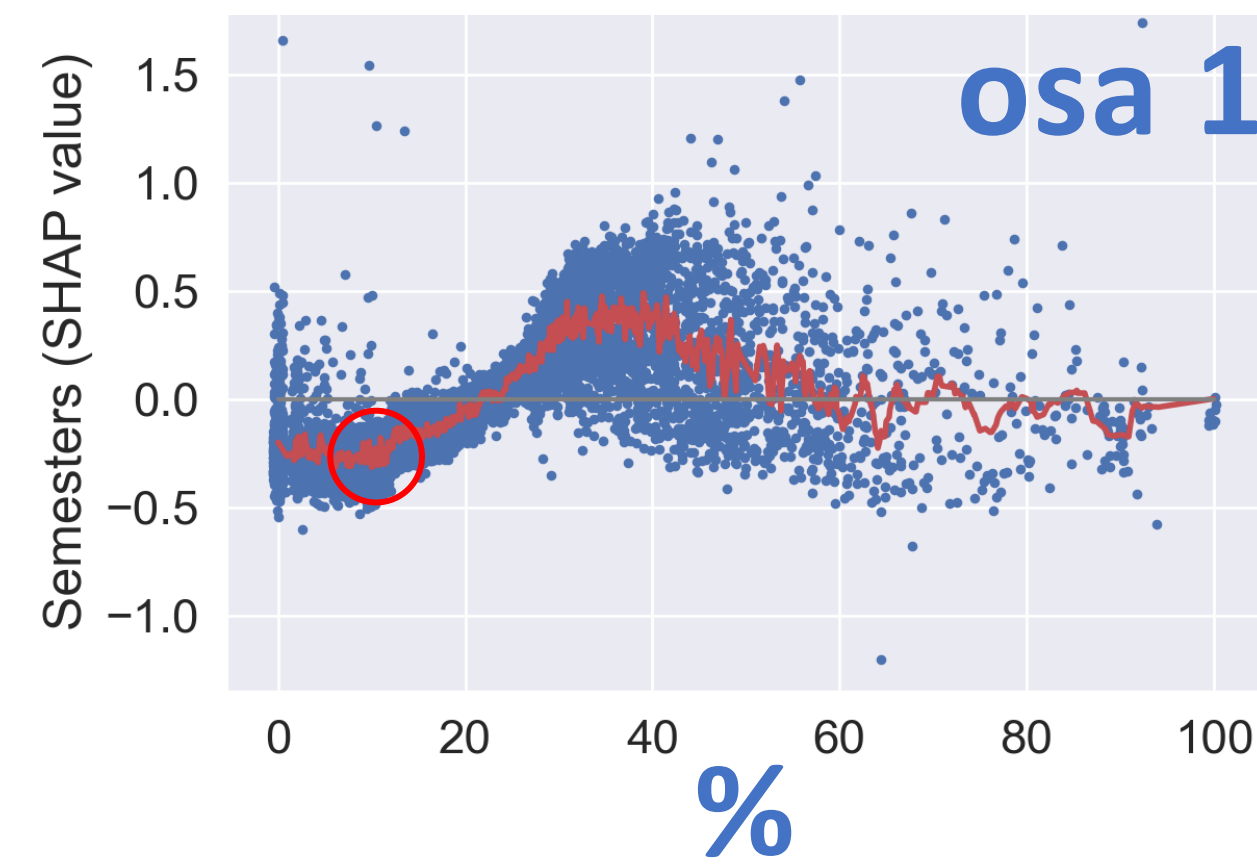
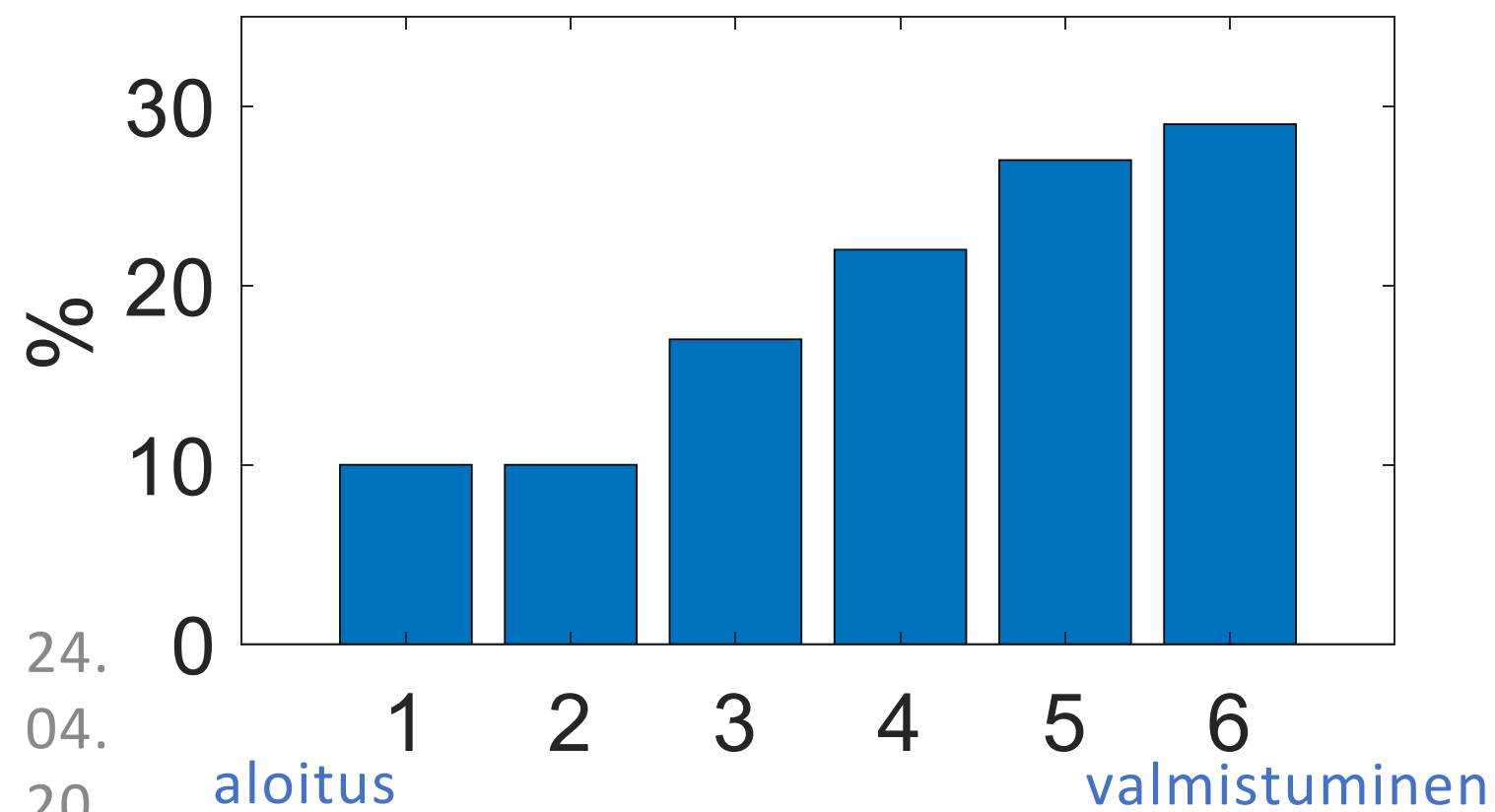
Tehtävien palautusaika (Moodle)

- Kategoriat (7 kpl) deadlineen mukaan:
 - myöhässä
 - Viimeiset 6h
 - Viimeiset 24h
 - 1-2 päivää
 - 3-6 päivää
 - 7-14 päivää
 - Yli 23 päivää
- Aina “viime tipassa” palauttaminen hidastava tekijä
- Huom: Paljon puuttuvaa dataa



Opintojen ajoituksen vaikutus

- Opiskeluaika jaettiin 6 osaan ja prosenttiosuudet kurssien jakautumiselle
- Esimerkki: 6v opiskelu → 1 osa = 1 vuosi
- Optimaalinen opintojen ajoitus olisi ”takapainotteinen”:

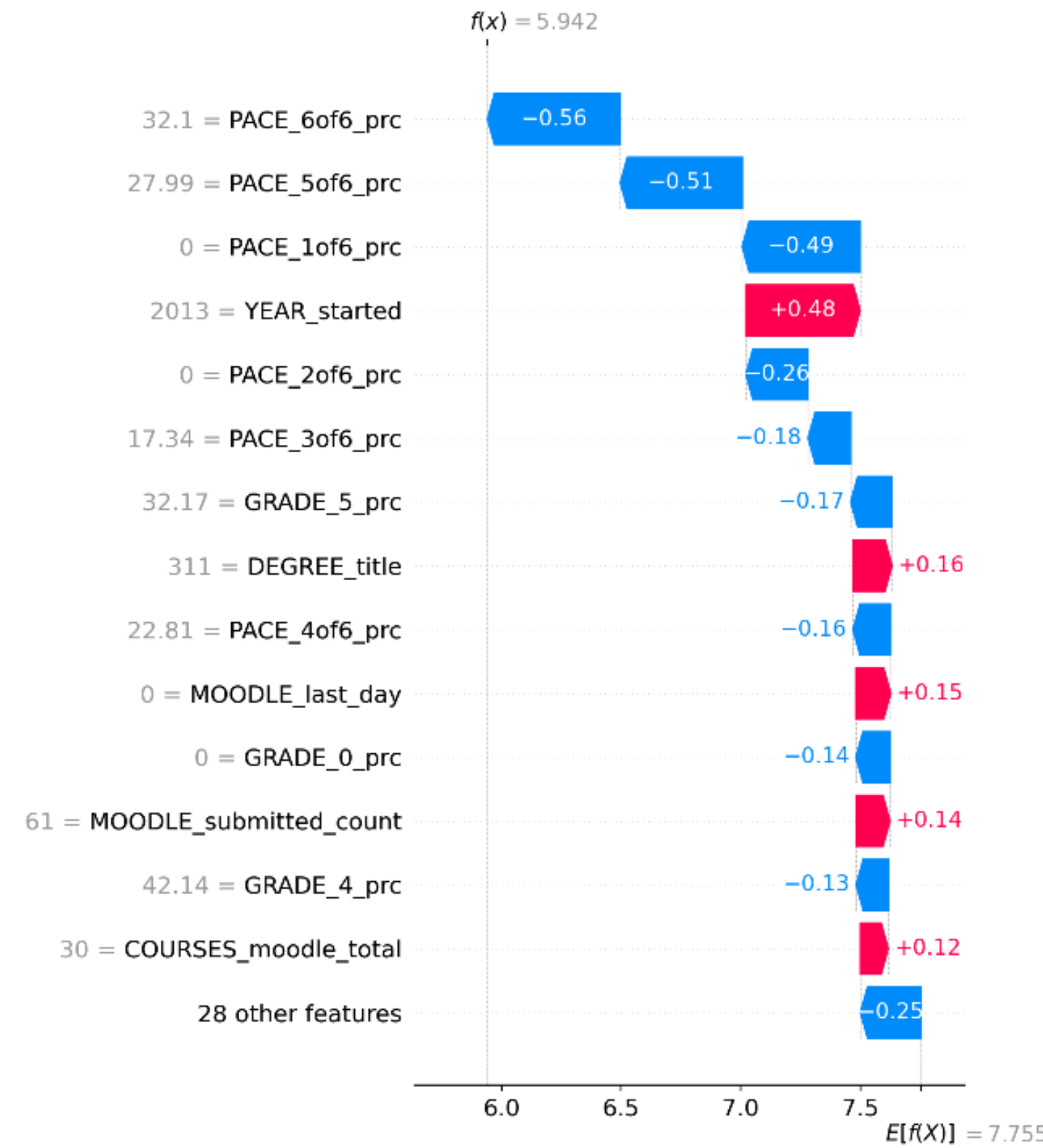


Korkeakoulujen IR-verkosto %

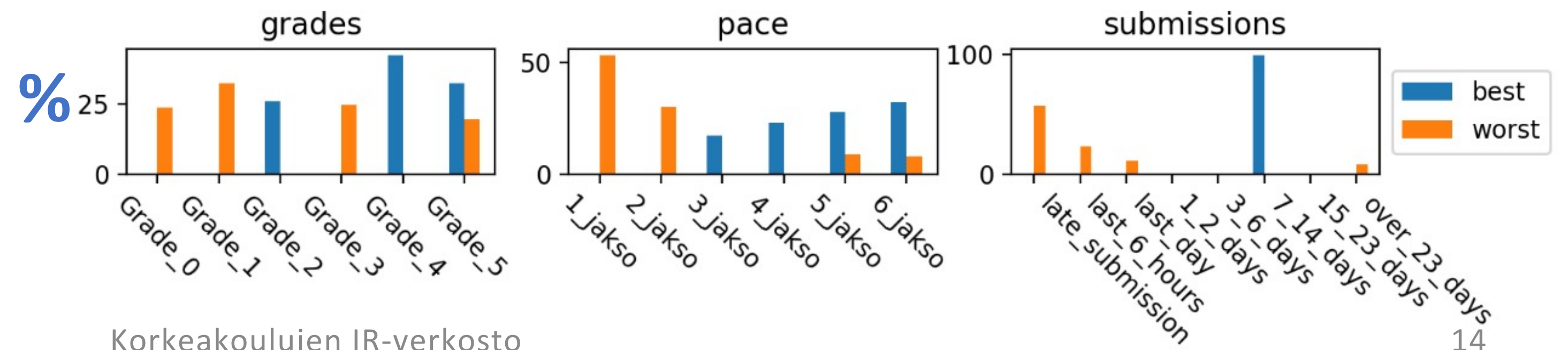
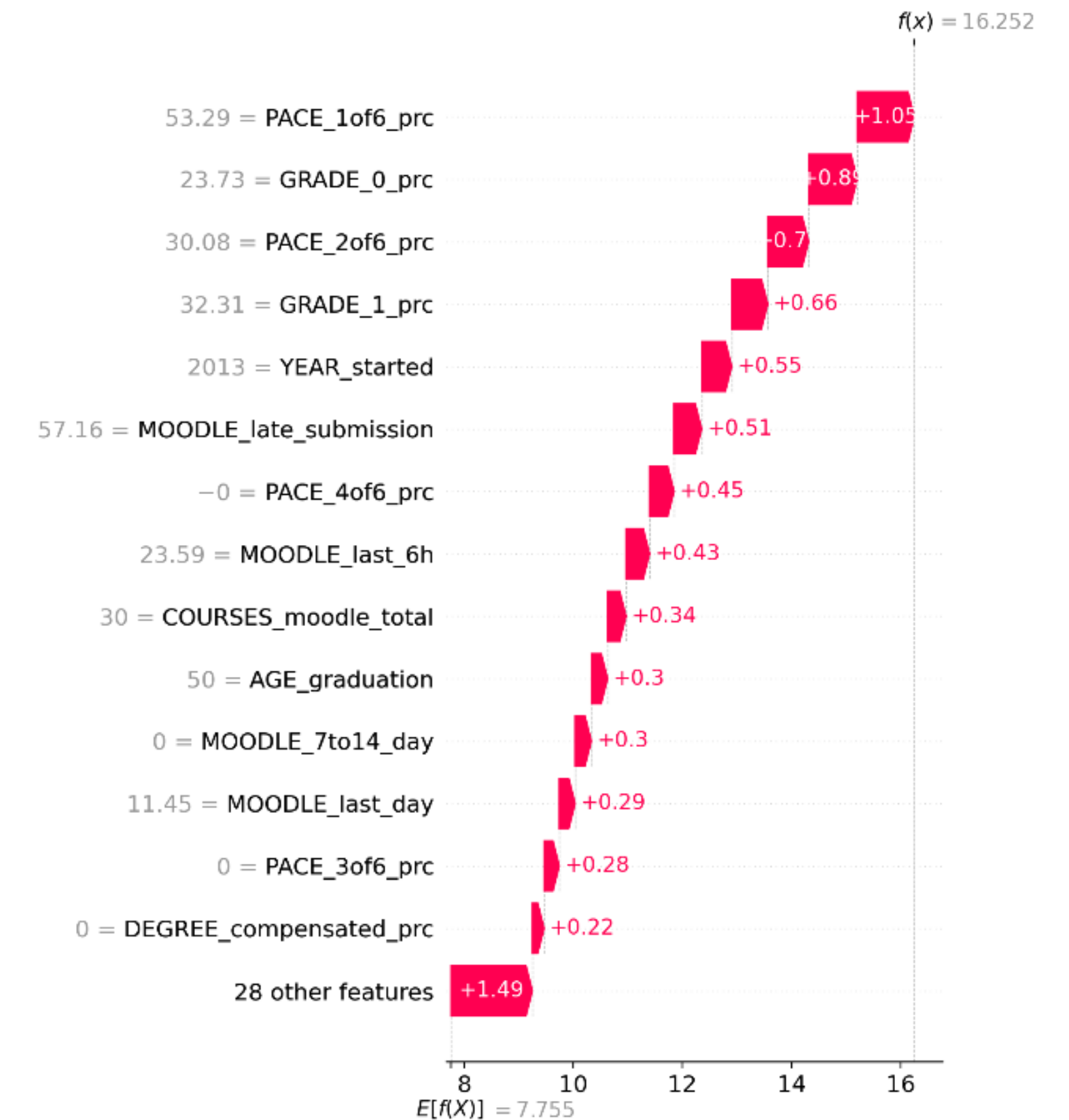
Opiskelutavan optimointi

- Käytetään mallia löytämään **optimaaliset parametrit** erilaisille tapauksille
- Tässä opiskelijana kuvitteellinen 50v nainen, joka on uudelleenkouluttautumassa
- Miten GRADES, PACE ja MOODLE valittava, jotta tulee **hitain** ja **nopein** opiskeluajan ennuste?

Nopea: ennuste = 6 lukukautta



Hidas: ennuste = 16 lukukautta



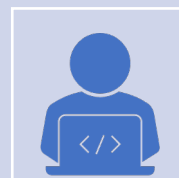
Analyysin yhteenveto

- Mallinnus kertoo eri osatekijöiden vaikutuksen opiskeluajan pituuteen
 - Opiskelijan oma tekeminen vs. kiinteät taustamuuttujat
- Nopeaa valmistumista edesauttavat mallin mukaan:
 - Opintojen kevyt aloitus
 - Hyvät arvosanat (mutta ei liikaa!)
 - Tehtävien palautus ajallaan
 - Pieni etäopiskelun määrä
- Mallissa ei mukana tärkeitä taustamuuttujia (esim. työssäkäynti, terveys, perhetilanne, asumistilanne jne.)
- Datat esikäsittely ja laatu ratkaisevaa (“garbage in, garbage out”)

Tulosten merkitys



Onnistuimme hyödyntämään Peppi- ja Moodle-dattaa koko opiskelijan opintopolulta ja rakentamaan yhteistyössä dataputken.



Ensimmäinen askel datan hyödyntämiseen:

Opiskelijan oma oppimisessa (reaaliaikainen)
Opinto-ohjaajan ja opettajan apuna (reaaliaikainen)
Oppimisen tietojärjestelmien, sisältöjen ja pedagogisten ratkaisujen apuna.



Jatkokehitys: Esim. Mobiilisovellus opiskelijalle, jossa ennakoiva malli ja liikennevalot



Kertoisi hyvissä ajoin, mikä on ennuste ja tarvittavat toimenpiteet jos ennuste ei miellytä.

The screenshot shows the Purdue Signals mobile application interface. At the top, it displays the AT&T logo, signal strength, and the time 8:49 PM. Below the 'Signals' header, there are three menu options: 'Detailed Report', 'Effort Tracker', and 'Help Resource'. The user's name 'Mary Major' and 'Fall Semester' are displayed. A table shows the student's performance in five courses across three interventions (Int 1, Int 2, Int 3). The courses are BIOL 101, GS 101, SPAN 310, STAT 303, and COM 150. Each course has a progress indicator (a circle) for each intervention. The SPAN 310 course shows a yellow circle for Int 2, indicating a warning or intervention point. The bottom of the screen features the Purdue University logo and a navigation bar with icons for back, forward, home, and a notification icon showing '3'.

Course	Int 1	Int 2	Int 3
BIOL 101	●	●	●
GS 101	●	●	●
SPAN 310	●	●	●
STAT 303	●	●	●
COM 150	●	●	●



Kysymysten ja keskustelun aika

Lisätietoja:

- Yliopettaja Lili Aunimo, lili.aunimo@haaga-helia.fi
- Vanhempi tutkija Janne Kauttonen, janne.kauttonen@haaga-helia.fi
- IT-asiantuntija Marko Vahtola, marko.vahtola@haaga-helia.fi