



VTT

Työpaketti 3:
**Tekoälypohjaisen
ennustemallin
kehittäminen**

**Jaakko Lähteenmäki
Emmi Antikainen**

13/09/2021 VTT – beyond the obvious

Ennustemalli

- Tavoitteena on kehittää koneoppimiseen perustuva ennustemalli:
 - käytettävissä olevien henkilötietojen perusteella ennustetaan henkilön joutumista paljon erilaisia sosiaali- ja terveystalvveluja tarvitsevien ryhmään tulevaisuudessa.
- Riskitodennäköisyyden lisäksi mallin tulee myös tuottaa tietoa riskiin vaikuttavista tekijöistä.
- Mallia tulee voida soveltaa sekä yksittäiseen henkilöön että henkilöryhmiin.
- Työn tuloksena tuotetaan tekoälymallin demonstraatioversio (Proof-of-Concept):
 - mahdollistaa konseptin esittelyn ja käyttäjäkokeilun (rajattu toiminnallisuus)
 - toteutus on paikallisesta tietojärjestelmäinfrastruktuurista riippumaton ja siten hyödynnettävissä myös kehitykseen osallistuvan sote-alueen ulkopuolella
 - operatiivinen käyttö päätöstukisovellusten kautta (eivät sisälly projektiin)

Ennustemalli

- Hyvinvointialueet
- Kansalliset toimijat



Päätöstuki-
sovellukset

Terveydenhoito

Palveluohjaus

Tietojohtaminen

...

Ennustemalli



Sote-alueen tietoallas,
kansallinen tietoallas
tai muu datalähde

Mallin kehittämisen vaiheet

- 1.* Data-aineiston, kohdejoukon ja muuttujien valinta sekä riskiryhmien määrittely**
- 2.* Aineiston käyttöön sisältyvät hakemukset, sopimukset ja järjestelyt**
3. Aineiston esikäsittely
4. Mallin koostaminen
5. Mallin validointi
- 6.* Käyttäjäkokeilu**

Tavoite:
3/2022 mennessä

* PHHYKYN osallistuminen keskeistä

Data-aineiston, kohdejoukon ja muuttujien valinta sekä riskiryhmien määrittely

- Kohdejoukko: 65-90 vuotiaat asukkaat
- Ajanjakso: tiedot viimeisen 5 vuoden ajalta, jos mahdollista
- Hyödynnettäviä henkilötietoja voisivat olla mm.
 - demografiatiedot
 - terveystietojen palvelutapahtumat (avohoidon käynnit ja sairaalajaksot) ja niihin liittyvät tiedot, kuten lääkitystiedot, diagnoosit, toimenpiteet ja DRG-ryhmittely (Diagnosis Related Group)
 - sosiaalihuollon palvelujen käyttö (ikäntyneiden laitoshoido, ympärivuorokautisen hoidon asumispalvelut, kotihoito, omaishoidon tuki, päihde- ja mielenterveyspalvelut)
- Seuraavat askeleet (yhteistyö PHHYKYn kanssa):
 - täsmennykset, mitä tietoja on käytettävissä ja mitkä ovat keskeisimmät tiedot
 - riskiryhmien kriteerien määrittely (ts. millä perusteella henkilö luokitellaan monialaisesti palveluja tarvitseviin)

Aineiston käyttöön sisältyvät hakemukset, sopimukset ja järjestelyt

- Kaksi etenemistapaa
 - omissa rekistereissa olevan tiedon käyttö tietojohdamiseen → ei tarvita Findatan tietolupaa
 - tieteellisin menetelmin toteutettava rekisteritutkimus → tarvitaan Findatan lupa
- Molemmissa tapauksissa PHHYKY:n tukea tarvitaan (mm. tietojen määrittely/poiminta ...)
- Tietojenkäsittelysopimus (DPA): PHHYKY / VTT
- Datan käyttöympäristö ja etäkäyttömahdollisuudet
- Seuraavat askeleet (yhteistyö PHHYKYn kanssa):
 - etenemistavan valinta (PHHYKY)
 - valmistelut käyntiin (hakemukset, sopimukset, järjestelyt)

Jatkotoimenpiteet

- tiedot käytettävissä olevasta data-aineistosta
- riskiryhmien kriteerit
- etenemistavan valinta aineiston käyttöön liittyen