

Kela|Fpa[®]

Tekoälykokeiluja



1kk kokeilu nykyisen chatbot-tuotteen LLM (large language model) -ominaisuuksilla 11/2023

- Tavoite:
 - Kokeilla miten LLM-ominaisuudet voisivat vähentää chattibotin ylläpitoon liittyvää manuaalista työtä ja tutkia miten LLM-ominaisuudet toimivat Kelan kontekstissa
- Havainnot:
 - LLM-ominaisuuksilla voisi automatisoida osan tällä hetkellä manuaalisesti tehtävistä chattibotin ylläpitotöistä, esim. sisällön tuottamisessa.
 - Tämä olisi edelleen linjassa digipalvelulain kanssa, sillä bottikuiskaaja (chattibotin kouluttaja ja sisällön ylläpitäjä) voisi hyväksyä tekoälyn tuotokset, kuten opetuslauseet ja vastausehdotukset, ennen kuin ne otetaan käyttöön.

1kk kokeilu nykyisen chatbot-tuotteen LLM (large language model) -ominaisuuksilla 11/2023

- Havainnot:
 - Kokeilimme myös kielimallia, joka oli opetettu kela.fi-sivujen materiaalilla. Tämä ei lyhyessä kokeilussa tuottanut hyviä tuloksia, sillä chattibotti ei osannut vastata kysymyksiin ja generoida kela.fi-sivuston materiaalin pohjalta hyviä vastauksia.
 - Totesimme, että kela.fi-sivusto on lähteenä liian suppea ja tiedon pitäisi olla eri muodossa, jotta kielimalli voisi käyttää sitä onnistuneesti. Kokeilu jäi vain kuukauden mittaiseksi, koska kyseessä oli chatbot-toimittajan tarjoama maksuton kuukauden kokeilu.
- Esteet:
 - Jotta voisimme hyödyntää nykyisen Boost.ai-tuotteen avustavia LLM-ominaisuuksia, tarvitsemme saas-luvan. Nykyinen chatbot-ratkaisu on konesalissa.

Selittäjä-kokeilu

- Kelassa kokeiltiin voisivatko toimeentulotuen tai työttömyysetuuden parissa työskentelevät ratkaisuasiantuntijat hyödyntää etuuskäsittelyssä tukena generatiivista tekoälyä, jolle on opetettu etuuden ratkaisutyön ohjeistus.
- Havainnot:
 - Vaikuttaa siltä, että tekoäly ei ole vielä kypsä tämän tyyppiseen käyttöön, sillä hallusinointia ilmeni, vaikka lähdemateriaali oli Kelan virallinen etuusohje
 - Tiedon luotettavuuden kannalta on olennaista, että tekoäly näyttää tarkalleen mistä ohjeistuksen kohdasta on vastauksensa muodostanut.
 - Jotta tekoällyn toiminnan saisi optimoitua, lähdemateriaalin pitäisi olla tietyllä tavalla koostettu rakenteiseen muotoon. Nykyinen etuusohje ei sellaisenaan tuota tarpeeksi hyviä tuloksia.

Asiakaspalautteiden analysointi

- Tarkoitus on käyttää laajaa kielimallia Kelan ja Kantan asiakaspalautteiden analysointiin.
 - Analysoinnilla tarkoitetaan tässä luokittelua, sävyn tunnistusta, tiivistelmien tekemistä palautteiden osajoukoista ja trendaavien teemojen tunnistamista. Data ei sinällään ole henkilödataa, koska se kerätään anonyymisti.
 - Ei voida täysin välttyä siltä, etteikö se sisältäisi henkilötietoja, koska palautteen kirjoittaja kirjoittaa, mitä kirjoittaa.
 - Sisältödatan ongelmilla kai tarkoitetaan sitä, että dataa jouduttaisiin siirtämään kielimallin ylläpitäjälle. Tämän riskin olemme poistaneet valitsemalla avoimia kielimalleja, jotka voidaan asentaa Kelan konesaliin ja näin tieto mallilta ei vuoda yhtään mihinkään.
 - Osumatarkkuuden ei tarvitse olla täydellinen, koska tulos on tarkoitettu helpottamaan asiantuntijan työtä. Nyt kaikki palautteet luetaan, mutta tällä luotaisiin yleiskuva ja pystyisi aina etsimään alkuperäiset palautteet esim tietyistä teemoista.
 - Tämä olisi tukityökalu ja siten riittävän hyvä osumatarkkuus riittää. Säästäisi joka tapauksessa valtavasti työaika.

Työdataekosysteemi –kokeilu

Työdataekosysteemi –kokeilu:

Kokeilun tavoitteena on:

- tuottaa yhdessä esimerkki, jonka pohjalta luomme sotudigivision toimeenpanolle konkreettisia askeleita
- nostaa kyvykkyyttämme toimia dataekosysteemeissä. Käyttötapausten kautta opitaan lisää ja voidaan jakaa oppeja, omalle ja muille organisaatioille.
- tuottaa yleistyksiä ja yleisiä suosituksia muihin elämäntapahtumiin ja dataekosysteemeihin.

Helsingin kaupungin työllisyyspalvelujen, kaupunginkanslian ja Kelan yhteinen kokeilu syksyllä 2024.

Fokuksessa on työtön/työkyvytön asiakas.

Kokeilun pohjana on molemmissa organisaatioissa tehdyt asioinnin ja tiedon liikkumisen ongelma-kartoitukset.

Mukana työllisyyspalveluista vastaavia, palvelumuotoilijoita, datatieteilijöitä, datainsinöörejä ja muita digitalisaation edistäjiä

Kokeilu toteutetaan kahdella radalla

Ennakollisesti parempi kuva asiakkaan työkyvyn riskeistä, joka auttaa tunnistamaan intervention paikat.

Kokeilussa aiomme

- selvittää, mitä Kelan ja Helsingin dataa tarvitaan ja miten tiedot ovat yhdistettävissä saadaksemme käsityksen asiakkaan työllistymisen esteistä/mahdollisuuksista
- tunnistaa yhdessä potentiaalisia työkyvyttömyyttä ennustavia muuttujia
- (määrittää sopivia aikaisia intervention paikkoja)

Yhteinen tilannekuva työttömästä asiakkaasta, jotta voimme ohjata asiakkaan oikeaan palveluun.

Kokeilussa aiomme

- selvittää, mitä Kelan ja Helsingin dataa tarvitaan ja miten tiedot ovat yhdistettävissä saadaksemme käsityksen asiakkaan työllistymisen esteistä/mahdollisuuksista
- Luoda (edes kömpelön) tilannekuvan työttömästä asiakkaasta virkailijan toimenpiteitä varten

Ennustemalli – lähtötilanne nyt Kelassa

- Kesällä 2024 Kelassa kehitetyn mallin avulla voidaan havaita "melko varmasti" tukien varaan jääminen palkansaajien osalta kahden vuoden päästä (tulorekisteridata apuna). Pystymme tuottamaan hyvän ennustemallin.
- Ennustemallia voidaan hyödyntää esim. signaaleina Kelan asiantuntijoille siten, että listataan suurimmassa riskissä olevat asiakkaat ja pyritään auttamaan heitä esim. kuntoutuksen keinoin.
- Mallin hyödyntämisen haasteena on se, että monen mallin tunnistaman asiakkaan tilanne on jo liian paha.

Ennustemalli – aiemmat kokeilut

Mallin kehittäjä	Eläketurvakeskus 2018	Kela 2021	Kela 2023
Tutkimuskysymys:	...ennustamaan työkyvyttömyyden alkamista kahden vuoden kuluttua?	...ennustamaan sp-raham saajien työkyvyttömyyseläkkeen riskiä?	...ennustamaan sp-raham saajien työkyvyttömyyseläkkeen riskiä?
"Tarkkuus*"	78 %	78 %	79 %
Aineisto	240 000 aiemmin työkyvyttömyyseläkkeelle joutuneen henkilön tietoista ajalta 2-10 vuotta ennen työkyvyttömyyseläkkeen alkamista	Sairauspäivärahaa saaneet Kelan asiakkaat tietyn ajanjakson sisällä	Sairauspäivärahaa saaneet Kelan asiakkaat tietyn ajanjakson sisällä

*Huomaa, että käsite "tarkkuus" on tässä esityksessä erittäin yksinkertaistettu kiteytys ennustemallin luotettavuudesta. Tulosten tarkempi analyysi vaatii data science –osaamista.

Ennustemalli – vuoden 2024 kokeilut:

Mallin
kehittäjä

Kela 2024 v1

Kela 2024 v2

Tutkimus-
kysymys

...ennustamaan tukien
varaam jäämisen riskiä?

...työkyvyttömyysriskin
kvantifiointiin sillä hetkellä, kun
asiakas tulee työllisyyspalveluiden
asiakkaaksi (riski seuraavan
kahden vuoden sisällä)

Tarkkuus

86 %

?? %

Aineisto

Sairauspäivärahaa saaneet Kelan
asiakkaat tietyn ajanjakson sisällä,
lisättyä tulorekisteritiedoilla

Työttömyysetuutta Kelasta hakeneet +
tulorekisteritiedoista ansiopäivärahan saajat

Uudet muuttajat Kelan aineistosta?

- Työvoimapolitiittisten lausuntojen sisältö (TE-palvelujen koodit)
- Ansiopäivärahan tiedot? (Kassan tietoja, rajoitetusti saatavilla Kelassa, tulorekisteristä saatavissa)
- Opintopisteiden kertyminen, opintojen keskeytyminen
- Hoitavan lääkärin työkykyarvio lääkärintodistuksen tekstisisällöstä? (nyt saadaan vain uusin diagnoosikoodi)
- Diagnoosien kokonaistilanne lääkärilausuntojen tekstisisällöstä?
- Kelan muut palvelut, joissa asiakkaat olleet (esim. MAP)?

Yhteisen datan ennustemalli vuodelle 2025?

Mallin
kehittäjä

Kela & Helsinki 2025

Tutkimus-
kysymys

...työkyvyttömyysriskin
kvantifiointi sillä hetkellä, kun
asiakas tulee työllisyyspalveluiden
asiakkaaksi (riski seuraavan
kahden vuoden sisällä)?

Tarkkuus

?? %

Aineisto

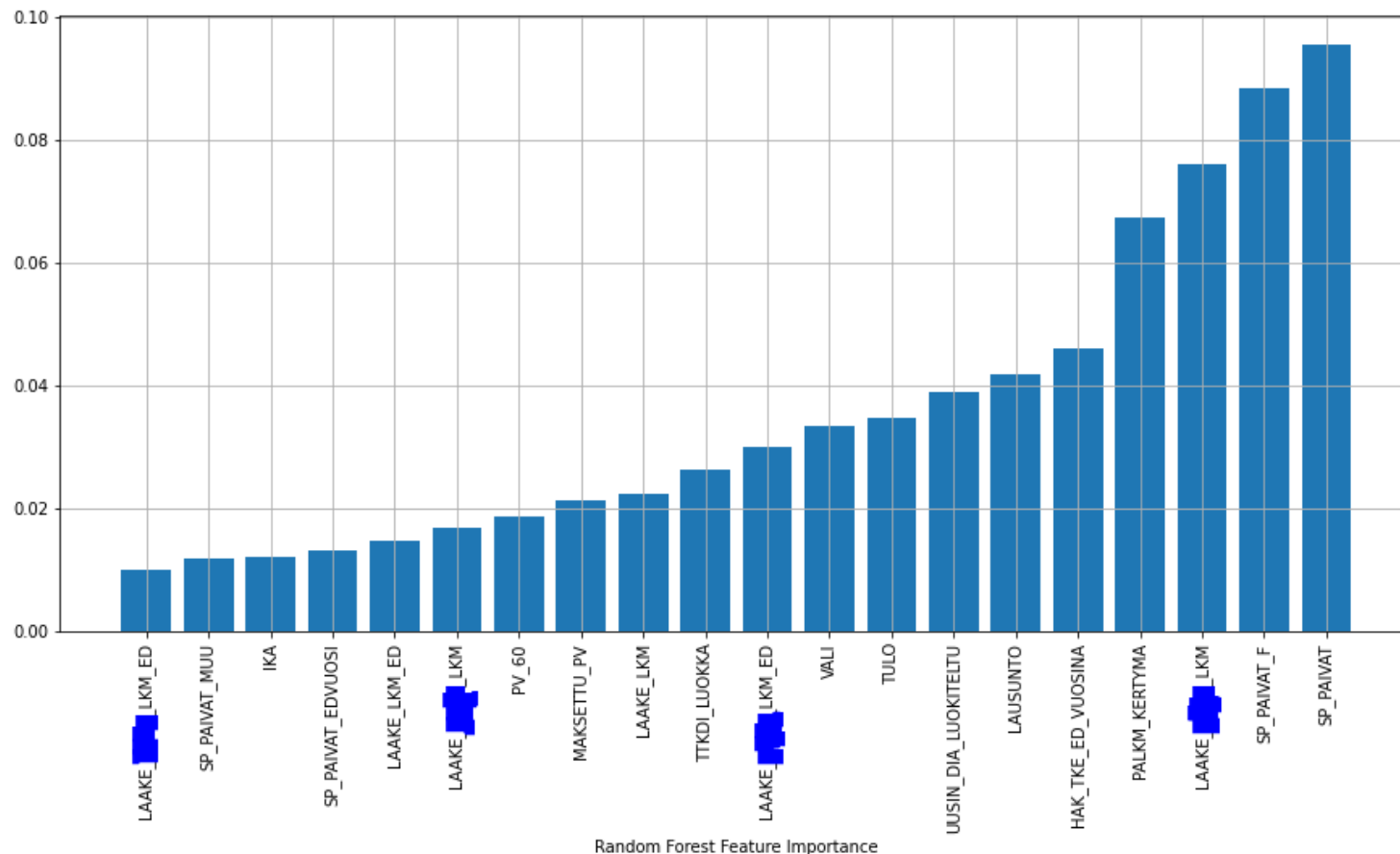
Kelan aiemman kokeilun (2024 v2) aineisto
ja Helsingin työllisyysdataa yhdistettynä

Uudet muuttajat Helsingin työllisyysdatasta?

- Työhistoriatiedot (esim. toistuvasti päättyvät työsuhteet)
- Koulutushistoriatiedot (Onko paljon keskeytyksiä, TUVA:t yms)
- Mihin palveluihin asiakas on ohjautunut Helsingissä (Työllisyyspalvelut: TYP yms)
- Asiakkaan oma arvio työkyvystään
- -----
- Yhteydenpito asiakkaan kanssa (Asiakas ei saavu sovittuihin tapaamisiin)
- Työllisyyspalveluiden asiakkuustiedot (Katkeaa esim. jos ei tule paikalle.)

Käytettyjen muuttujien "tärkeysjärjestys"

- Muuttujia 122 kpl
- Viereisessä kuvassa top 20 tärkeintä muuttujaa
- Mm. sairauspäivärahopäivät ja lääkeostot korostuvat



Muuttajat selkokielellä

1. Sairauspäivärahopäivien lkm
2. Sairauspäivärahopäivien lkm F-diagnoosilla
3. Lääkeostokertojen lukumäärä tietyllä ATC-koodilla
4. (kertynyt pa-lukumäärä?)
5. Hakenut tk-eläkettä edellisten 2 vuoden aikana vähintään kerran
6. Sairauspäivärahan hakemuksessa käytetyn lausunnon tyyppi (A- tai B-lausunto)
7. Viimeisin diagnoosikoodi, jolla sp-rahaa on haettu (vain osatotuus, lausunnoissa voi olla lisää d-koodeja)
8. Tulorekisterin tulotiedot - sairauspäivärahan laskennan perusteena käytetty tulo
9. Väli (?) - Maksun ja työkyvyttömyyden alun välisten päivien lukumäärä
10. Lääkeostokertojen lukumäärä edelliseltä vuodelta tietyllä ATC-koodilla
11. Työ- / toimintakoodi työkyvyttömyyden alkaessa (Helsingillä parempi/tarkempi?)
12. Lääkeostokertojen lukumäärä yhteensä
13. Maksettujen sp-rahojen määrä euroissa
14. Onko 60 sp-rahopäivää tullut täyteen (0/1)?
15. Lääkeostokertojen XXX? lukumäärä
16. Lääkeostokertojen lukumäärä edellisenä vuonna
17. Sairauspäivärahopäivien lkm edellisenä vuonna
18. Ikä vuoden tarkkuudella
19. Sairauspäivärahopäivien lkm (muut??)
20. Lääkeostokertojen lukumäärä edelliseltä vuodelta tietyllä ????

Lainsäädännön näkökulma: TYM-laki 2025 alkaen

- ”Työttömän suostumuksesta riippumatta työvoimaviranomaisella, hyvinvointialueella ja Kansaneläkelaitoksella on oikeus salassapitosäännösten ja muiden tiedon saantia koskevien rajoitusten estämättä saada toisiltaan sekä tässä laissa tarkoitettujen tehtävien hoitamiseksi käyttää 11 §:n 2 momentissa tarkoitettuja työtöntä koskevia **tietoja, jotka ovat välttämättömiä** työllistymistä edistävän monialaisen tuen yhteistoimintamallin järjestämiseksi työttömälle”.
- Mitkä palvelut ja asiat kuuluvat ”yhteistoimintamallin järjestämiseen”?
- Mitkä tiedot ovat **välttämättömiä**?
- Helsingin työllisyyspalvelut panostavat ennakkolliseen toimintaan ja haluavat tunnistaa vaikuttavat ja varhaiset intervention paikat:
 - Onko yhteinen ennustemalli palvelu, joka kuuluu yhteistoimintamalliin? Voidaanko tällä perusteella saada ja hyödyntää toistemme tietoja asiakkaista?

Asiantuntijoita Kelassa

Janne Mattila, Veera Hyytiä, Päivi Rajaharju

Kela|Fpa[®]