

Mitä voimme oppia ihmisälystä tekoälyn avulla?

Osa 1: Logiikka ihmisälyn mallina.

Psykologinen oletus:

Ihmiset ovat pääpiireissään rationaalisia.

Logiikka kertoo, mitä rationaalisuus on.

Joten: Ihmisten toiminnan täytyy perustua logiikalle.

Psykologinen oletus:

Ihmiset ovat pääpiireissään rationaalisia.

Logiikka kertoo, mitä rationaalisuus on.

Joten: Ihmisten toiminnan täytyy perustua logiikalle.

Tietokone mielen mallina:

1. Logiikan kehitys matematisoiduksi teoriaksi.
2. Logiikan formalisointi mekaaniseksi laskennoksi.
3. Tietokoneiden synty matemaattisesta logiikasta.
4. Tietokoneet mekaanisina järkeilykoneina.
5. Tietokoneohjelmat järjenkäytön psykologisina malleina.

Ruusut tarvitsevat vettä.

Kasvit tarvitsevat vettä.

siis: Ruusut ovat kasveja.

Ruusut tarvitsevat vettä.

Kasvit tarvitsevat vettä.

siis: Ruusut ovat kasveja.

Käärmeet ovat vaarallisia.

Aseet ovat vaarallisia.

siis: Käärmeet ovat aseita.

1. Matti katsoo Liisaa, Liisa katsoo Perttiä.
2. Matti on naimisissa, Pertti ei ole naimisissa.

Katsooko joku, joka on naimisissa, jotakuta, joka ei ole?

Kyllä / Ei / Ei voi päätellä

1. Matti katsoo Liisaa, Liisa katsoo Perttiä.
2. Matti on naimisissa, Pertti ei ole naimisissa.

Katsooko joku, joka on naimisissa, jotakuta, joka ei ole?

Kyllä / Ei / Ei voi päätellä

Jos Liisa **ei ole** naimisissa: Kyllä (Matti → Liisa)

Jos Liisa **on** naimisissa: Kyllä (Liisa → Pertti)

Logiikka ihmisjärjen mallina

Sisältö ja tuttuus, **ei looginen muoto**, vaikuttaa ymmärretäänkö ongelma.

Emme yleensä järkeile loogisesti, vaikka ongelma olisi helppo ymmärtää.
Samat ongelmat toistuvat logiikan lisäksi muissakin järkeilytehtävissä.

Logiikka ihmisjärjen mallina

Sisältö ja tuttuus, **ei looginen muoto**, vaikuttaa ymmärretäänkö ongelma.

Emme yleensä järkeile loogisesti, vaikka ongelma olisi helppo ymmärtää.
Samat ongelmat toistuvat logiikan lisäksi muissakin järkeilytehtävissä.

Ihmiset käyttävät aikansa rationalisointiin, ei kriittiseen ajatteluun.

Rationalisointi: Järkevän kuuloisten perustelujen keksiminen lennosta.

Esim. Nisbett & Wilson 1977: Ostoskeskuskoe
Strandberg ym. 2018: Poliittiset mielipiteet.

Tärkeä tekoälytutkimuksen opetus jo alkuvuosilta oli:

- Arkinen todellisuus on erittäin monimutkainen.
- Silti ihmiset suoriutuvat arjestaan varsin sujuvasti.

(Esim.: Tulen elokuviin, jos saan illaksi vapaata.)

Tärkeä tekoälytutkimuksen opetus jo alkuvuosilta oli:

- Arkinen todellisuus on erittäin monimutkainen.
- Silti ihmiset suoriutuvat arjestaan varsin sujuvasti.

(Esim.: Tulen elokuviin, jos saan illaksi vapaata.)

Ogelma: Vaihtoehtoisten järkelynmallien puute.

Esim. Tieteenfilosofien käsitys kausaalisesta, induktiivisesta, jne. päättelystä.

Mitä älykkyys on?

Tärkeä tekoälytutkimuksen opetus jo alkuvuosilta oli:

- Arkinen todellisuus on erittäin monimutkainen.
- Silti ihmiset suoriutuvat arjestaan varsin sujuvasti.

(Esim.: Tulen elokuviin, jos saan illaksi vapaata.)

Ogelma: Vaihtoehtoisten järkelynmallien puute.

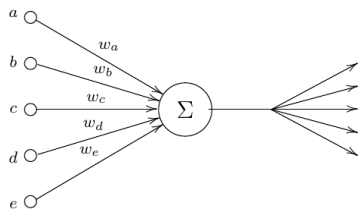
Esim. Tieteenfilosofien käsitys kausaalisesta, induktiivisesta, jne. päättelystä.

Mitä älykkyys on? **Kykyä toimia tarkoituksenmukaisesti monimutkaisissa ja vain osittain ennustettavissa ympäristöissä.**

Mitä voimme oppia ihmisälystä tekoälyn avulla?

Osa 2: Neuroverkot kognition mallina.

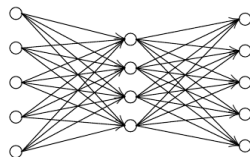
Neuroverkot aivotoiminnan malleina



tulevat yhteydet

summaatio

lähtevät yhteydet

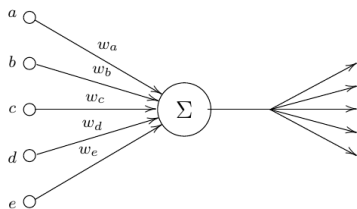


syötekerros

piilokerros

tulostekerros

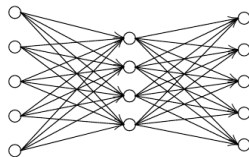
Neuroverkot aivotoiminnan malleina



tulevat yhteydet

summaatio

lähtevät yhteydet



syötekerros

piilokerros

tulostekerros

Signaalin ja vasteen assosiatiiivinen oppiminen.

Ajattelu = Signaali pyörii aivoissa?

→ Assosiatiiivinen psykologia.

Yhdelläkin tulostesolulla voidaan tehdä yllättävän monia asioita, kun se kytketään isoon syötekerrokseen ja opetetaan oikealla tavalla.

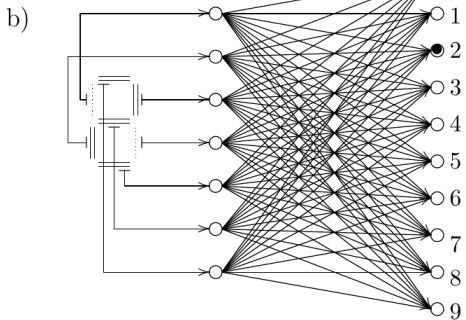
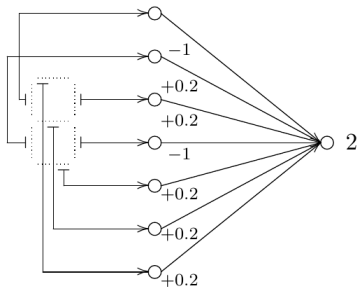
Yksinkertaisten hahmojen (esim. numeroiden) tunnistaminen,

Yhdelläkin tulostesolulla voidaan tehdä yllättävän monia asioita, kun se kytketään isoon syötekerrokseen ja opetetaan oikealla tavalla.

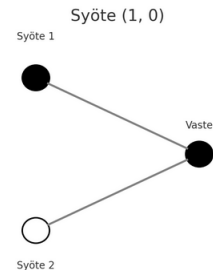
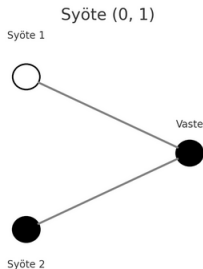
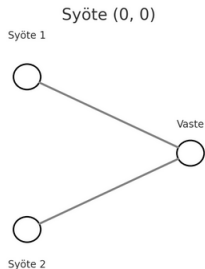
Yksinkertaisten hahmojen (esim. numeroiden) tunnistaminen,

Toisaalta kaikki edes kaikki loogiset operaatiot eivät onnistu, Esim. eksklusiivinen disjunktio.

Neuroverkot kognition mallina



Neuroverkot kognition mallina



XOR (Eri)

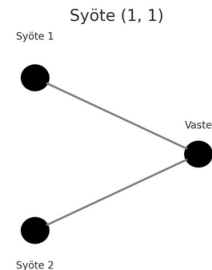
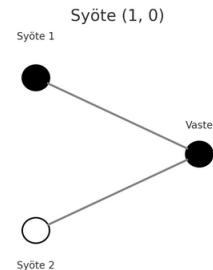
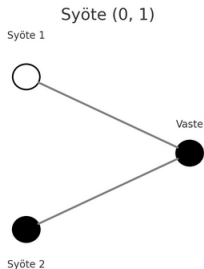
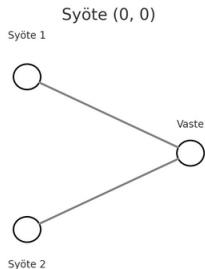
$$0 + 0 \mapsto 0$$

$$0 + 1 \mapsto 1$$

$$1 + 0 \mapsto 1$$

$$1 + 1 \mapsto 0$$

Neuroverkot kognition mallina



XOR (Eri)

$$0 + 0 \mapsto 0$$

$$0 + 1 \mapsto 1$$

$$1 + 0 \mapsto 1$$

$$1 + 1 \mapsto 0$$

Kytkemällä useampia kerroksia rajoitteet voidaan ylittää.

1960-luvulla ei osattu opettaa verkkoja, joissa on monta kerrosta
1980-luvulla keksittiin tähän opetusarlogritmi (backpropagation)

Kytkemällä useampia kerroksia rajoitteet voidaan ylittää.

1960-luvulla ei osattu opettaa verkkoja, joissa on monta kerrosta
1980-luvulla keksittiin tähän opetusarlogritmi (backpropagation)

Teoriassa verkoilla voi tehdä vaikka mitä.

Käytännössä laskentateho ja opetusaika muodosti ongelman.

Yksinkertaisilla verkoilla ei voitu tehdä kovin monimutkaisia asioita

Syväoppimisverkkojen läpimurto 2010-luvulla

Yksinkertaisesti laskentatehon kasvu, grafiikkaprosessorien käyttö, jne.
Opetusdatan huima kasvu. Data myös saatiin uusista lähteistä,
→ tekstit ja ihmisten keskustelut ja muu toiminta netissä jne.

Syväoppiminen: Kerroksilla voi olla reilusti syvyyttä (piilokerroksia).
Suuret kielimallit: Valtavasti soluja ja kytkentöjä, Valtavasti opetusdataa

Uusia opetusmenetelmiä: Opettamisen automatisointi, vahvistusoppiminen
Uusia verkkorakenteita: Konvoluutioverkot, transformer-arkkitehtuuri, jne.

→ Vallankumous suorituskyvyssä

Syväoppimisverkkojen läpimurto 2010-luvulla

Yksinkertaisesti laskentatehon kasvu, grafiikkaprosessorien käyttö, jne.
Opetusdatan huima kasvu. Data myös saatiin uusista lähteistä,
→ tekstit ja ihmisten keskustelut ja muu toiminta netissä jne.

Syväoppiminen: Kerroksilla voi olla reilusti syvyyttä (piilokerroksia).
Suuret kielimallit: Valtavasti soluja ja kytkentöjä, Valtavasti opetusdataa

Uusia opetusmenetelmiä: Opettamisen automatisointi, vahvistusoppiminen
Uusia verkkorakenteita: Konvoluutioverkot, transformer-arkkitehtuuri, jne.

→ Vallankumous suorituskyvyssä

HUOM: Sama laskentatehon kasvu olisi käytössä myös logiikamalleille.
Siispä **pelkkä** laskentateho ei selitä suorituskyvyn kasvua.

Mitä voimme oppia ihmisälystä tekoälyn avulla?

Osa 3: Tekoäly ihmisällyn tieteellisenä mallina

1. Uudet tekoälymallit eivät ole kehitetty mielen tai aivojen malleiksi.
2. Tarpeeksi tehokas oppimissysteemi oppii matkimaan mitä vain.
3. Uudetkin mallit ovat lopulta tyhmiä ja heijastavat vain oppimaansa.

1. Uudet tekoälymallit eivät ole kehitetty mielen tai aivojen malleiksi.
2. Tarpeeksi tehokas oppimissysteemi oppii matkimaan mitä vain.
3. Uudetkin mallit ovat lopulta tyhmiä ja heijastavat vain oppimaansa.

toisaalta taas:

2. On teoreettisia argumentteja, että assosiatiivinen äly on puutteellinen
→ **Näytöt** ovat kuitenkin näiden mallien puolella.
3. Eroaako tämä ihmisistä?

Tieteellisistä malleista

Simulaatiomallit: Esim. fysiikassa ja insinööritieteissä

Tieteellisistä malleista

Simulaatiomallit: Esim. fysiikassa ja insinööritieteissä

Teoreettiset mallit: Esim. Taloustiede ja yhteiskuntatieteet

- Idealisoituja ja yksinkertaistavia
- Ei mallinneta todellisia mikrotason mekanismeja
- Ennustusvoima voi olla heikkoa

Eivät mallinna lakeja vaan hypoteettisia säännönmukaisuuksia.
Tästä huolimatta ne voivat olla tieteellisesti hyödyllisiä.

Simulaatiomallit: Esim. fysiikassa ja insinööritieteissä
Teoreettiset mallit: Esim. Taloustiede ja yhteiskuntatieteet

- Idealisoituja ja yksinkertaistavia
- Ei mallinneta todellisia mikrotason mekanismeja
- Ennustusvoima voi olla heikkoa

Eivät mallinna lakeja vaan hypoteettisia säännönmukaisuuksia.
Tästä huolimatta ne voivat olla tieteellisesti hyödyllisiä.

- Selvittävät mahdollisia mekanismeja
- Auttavat tutkimaan niitä ja luomaan teorioita
- Hypoteettiset mekanismit voivat selittää todellisia ilmiöitä

- Yksinkertaistaminen on hyve, kun tavoite on ymmärtää.
- Ei välttämätöntä olla kohteen **mekanismien tarkka kuvaus**.
- Mallien tulisi **jotenkin laadullisesti käyttäytyä** kuten kohteensa.

- Yksinkertaistaminen on hyve, kun tavoite on ymmärtää.
- Ei välttämätöntä olla kohteen **mekanismien tarkka kuvaus**.
- Mallien tulisi **jotenkin laadullisesti käyttäytyä** kuten kohteensa.

Teoreettisten mallien tarkoitus on mahdollistaa mielenkiintoisten päätelmien tekemisen kohteestaan.

Siis:

1. Ei pidä paikkaansa, että logiikkamallit eivät kertoneet mielestä mitään, koska ne eivät olleet hyviä kognition malleja.
2. Ei pidä paikkaansa, että koneoppimismallit eivät kerro mielestä mitään, koska ne eivät edes yritä olla kognition malleja.

Mikä suurten kielimallien yms. opetus voisi olla?

1. Järkeily ei edellytä formaalia tai tietoista järkeilykykyä.
 - Tämä ei yllätä: arkinen todellisuus on epämääräinen ja monimutkainen.
2. Kielimallit yhdistelevät opetusdataa ja selittelevät mitä sattuu.
 - Nykyisten kielimallien typeryyskin lienee ihmismielen kuva.

Mikä suurten kielimallien yms. opetus voisi olla?

1. Järkeily ei edellytä formaalia tai tietoista järkeilykykyä.
 - Tämä ei yllätä: arkinen todellisuus on epämääräinen ja monimutkainen.
2. Kielimallit yhdistelevät opetusdataa ja selittelevät mitä sattuu.
 - Nykyisten kielimallien typeryyskin lienee ihmismielen kuva.
3. Inhimillinen rationaalisuus vaikuttaa olevan käytännöllinen kysymys:
 - Kykyä soveltaa ja yhdistellä tuttuja asioita uusissa tilanteissa.
 - Kyvyn voi parhaiten oppia jäljittelemällä ympäristön säännönmukaisuuksia.
 - **Ei tarvitse mallintaa mielensisäisiä prosesseja vaan vuorovaikutusta.**

Mikä suurten kielimallien yms. opetus voisi olla?

1. Järkeily ei edellytä formaalia tai tietoista järkeilykykyä.
 - Tämä ei yllätä: arkinen todellisuus on epämääräinen ja monimutkainen.
2. Kielimallit yhdistelevät opetusdataa ja selittelevät mitä sattuu.
 - Nykyisten kielimallien typeryyskin lienee ihmismielen kuva.
3. Inhimillinen rationaalisuus vaikuttaa olevan käytännöllinen kysymys:
 - Kykyä soveltaa ja yhdistellä tuttuja asioita uusissa tilanteissa.
 - Kyvyn voi parhaiten oppia jäljittelemällä ympäristön säännönmukaisuuksia.
 - **Ei tarvitse mallintaa mielensisäisiä prosesseja vaan vuorovaikutusta.**

(Kielimallit eivät opi inhimillisellä tavalla, mutta tilanne voi pian muuttua.)