

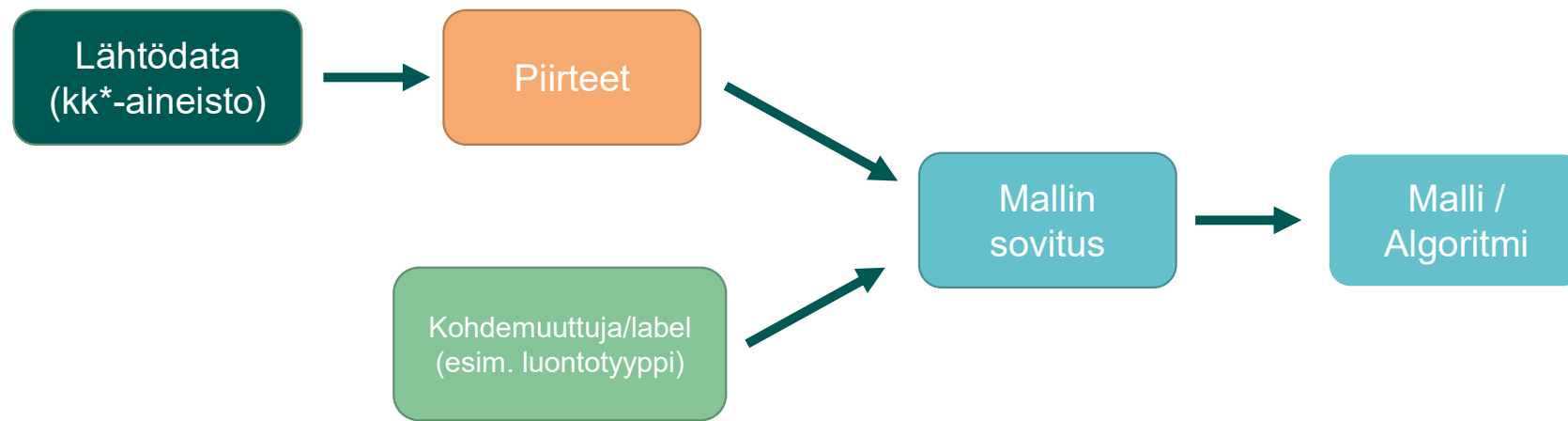
Koneoppimismallit tulkinnassa

Mikko Impiö, Syke

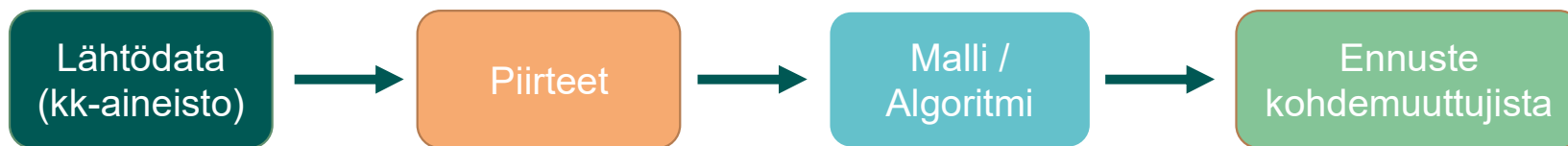


Suomen ympäristökeskus
Finlands miljöcentral
Finnish Environment Institute

Opetus



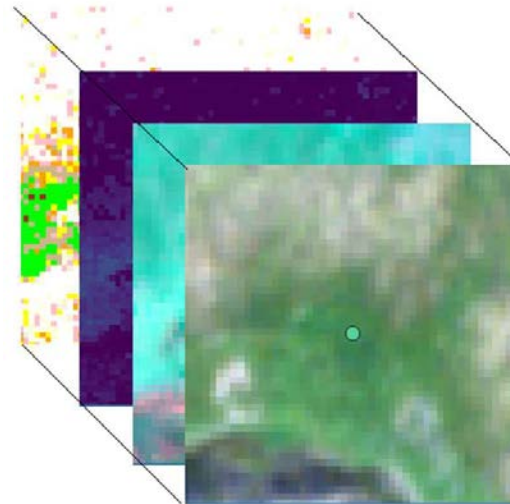
Tulkinta



Aineistot >> Algoritmit

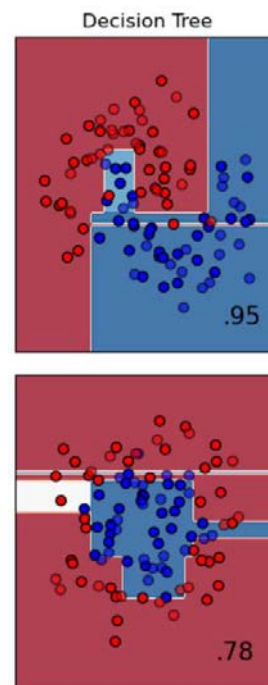
Pikselipohjaiset ja objektipohjaiset tulkintamenetelmät

- Pikselipohjaiset
 - AutoML-menetelmät
 - Satunnaismetsät (Random Forest)
- Objektipohjaiset
 - Syvät neuroverkot (CNN)

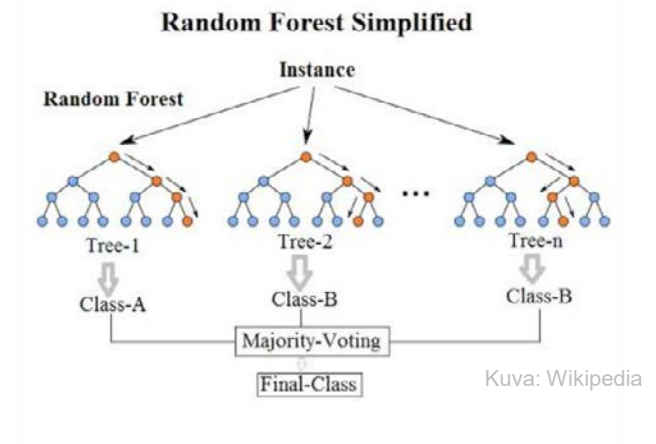


Satunnaismetsät (Random Forests)

- Perustuu usean päätöspuun konsensusukseen kohdemuuttujasta
- Yleisesti käytetty pikselipohjaisessa tulkinnaissa
- Selitettävä, intuitiivinen malli



Kuva: Sklearn /
Gaël Varoquaux



AutoML

- TPOT
 - AutoML-ohjelmakirjasto joka perustuu geneettiseen ohjelmointiin
 - Pystyy käymään läpi tuhansia eri mallikonfiguraatioita ja esiprosessointimenetelmiä
- AutoML-menetelmien käyttö helpottaa mallien löytämistä, mutta ei poista piirteiden valinnan tärkeyttä
- Myös muita samalla periaatteella toimivia kirjastoja (mm. Auto-Sklearn, AutoKeras)



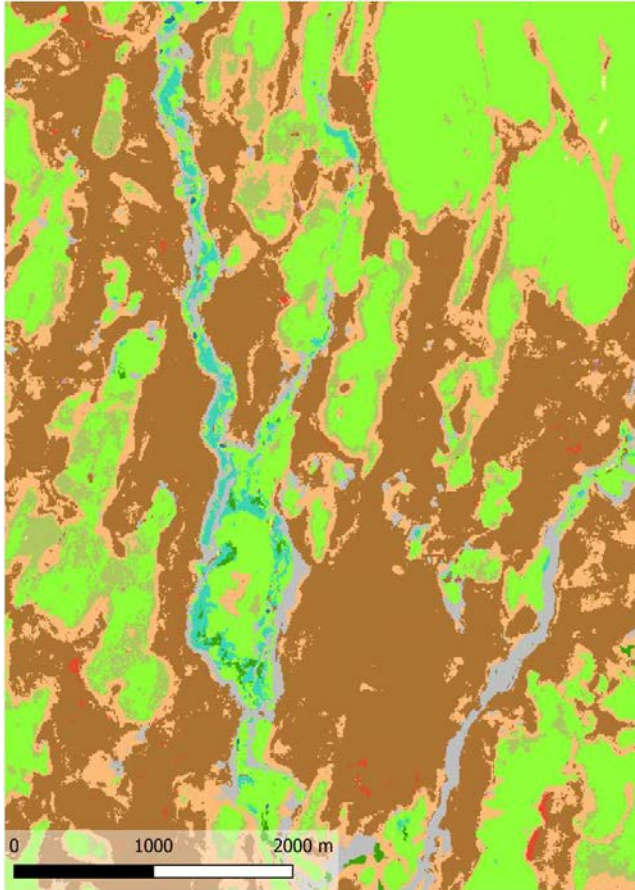
Syväoppiminen ja neuroverkot

- Käyttää tulkinnessa pikseliä ympäröivää aluetta
- Laskennallisesti raskaita, vaativat paljon dataa
- Mallien opetuksessa voidaan käyttää laajempia aineistoja
- Paljon potentiaalia jatkotutkimukselle ja kehitykselle

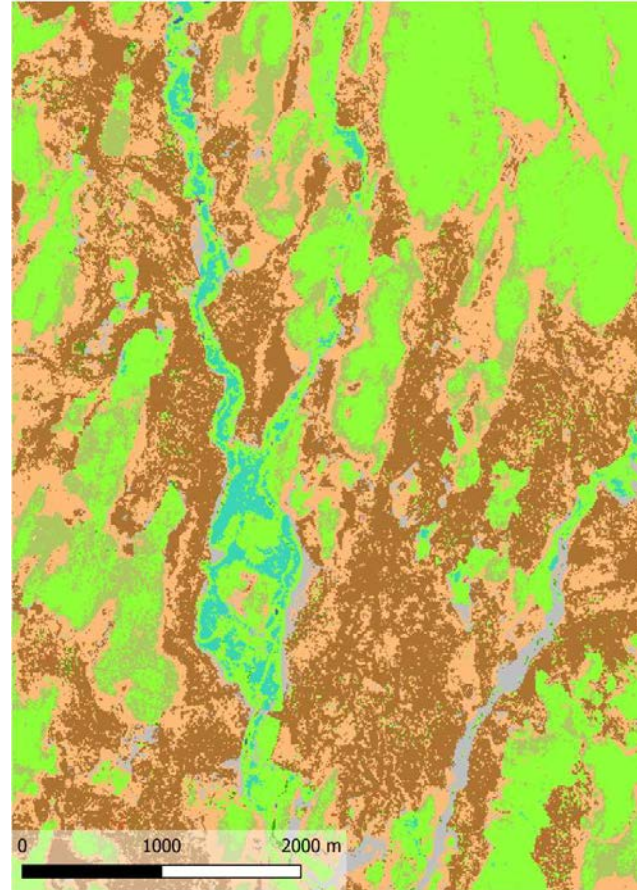


Opetusmateriaali,
josta
kohdemuuttuja
tiedossa

Potentiaalinen
opetusmateriaali



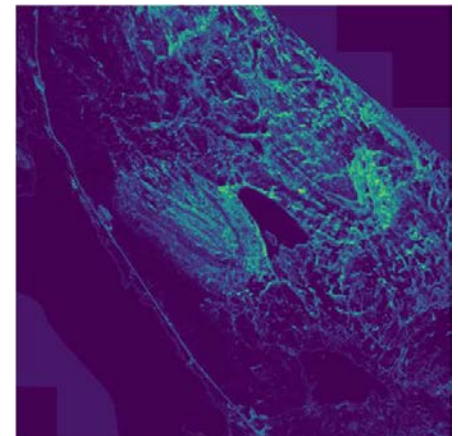
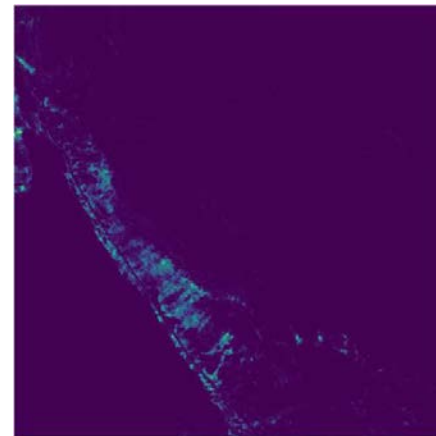
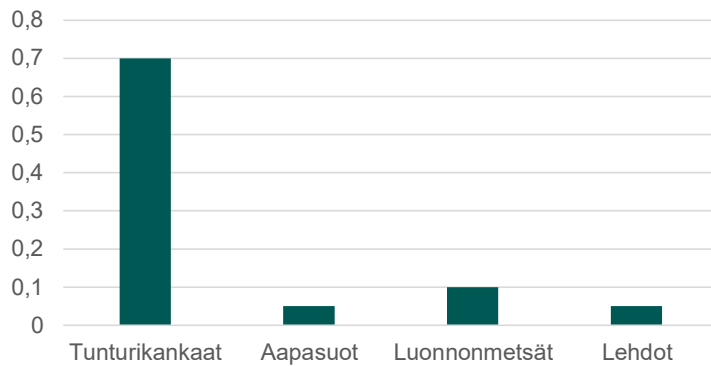
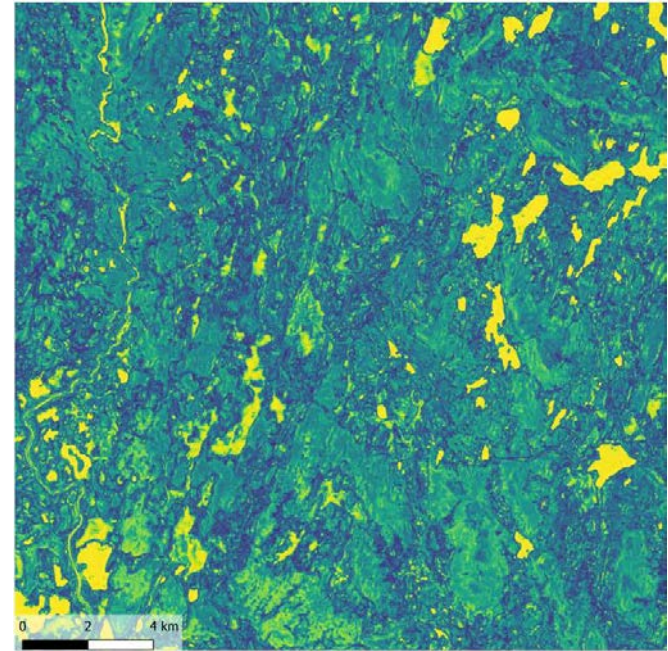
Neuroverkko



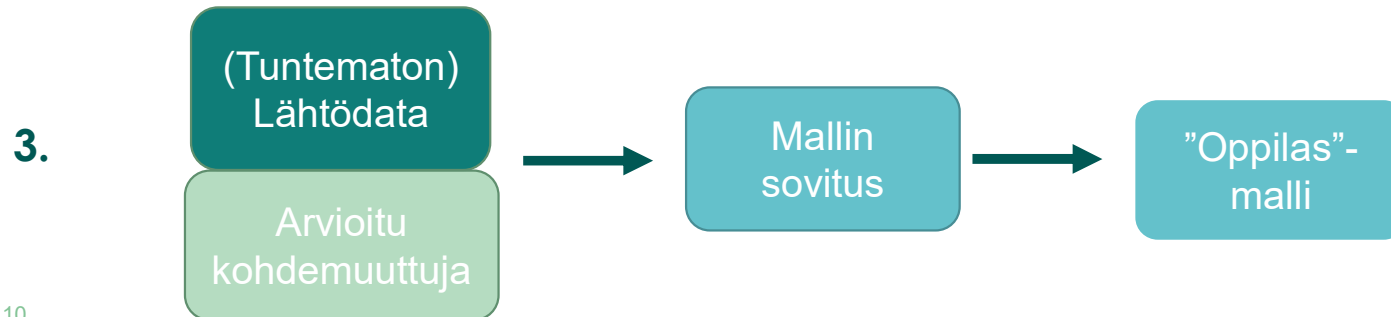
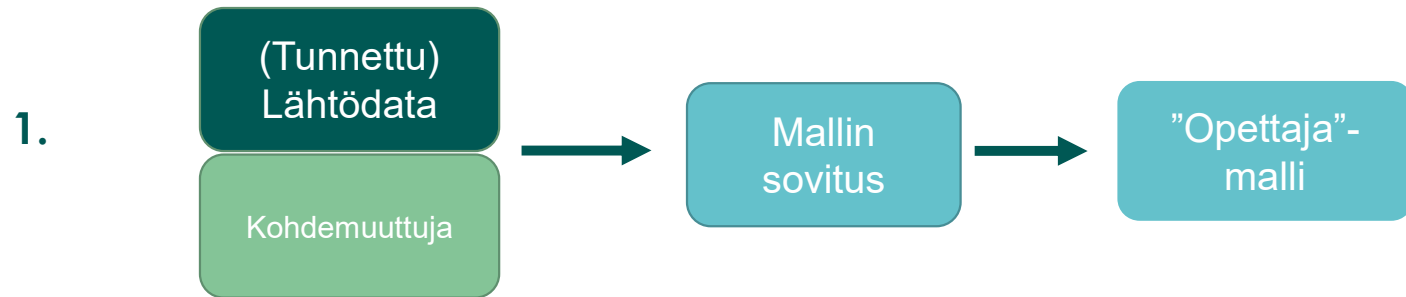
Random Forest

Luokitteluvarmuus

- Mallista riippuen on mahdollista laskea luokkakohtaisia ”varmuuksia” luokittelutulokselle
- Luokkakohtaisten varmuuksien summa = 1
- Maksimivarmuuden luokka valitaan (yleensä) lopulliseksi luokaksi
- Varmuuskartat (kuva) osoittavat epävarmat luokittelualueet



Puoliohjattu oppiminen





Suomen ympäristökeskus
Finlands miljöcentral
Finnish Environment Institute