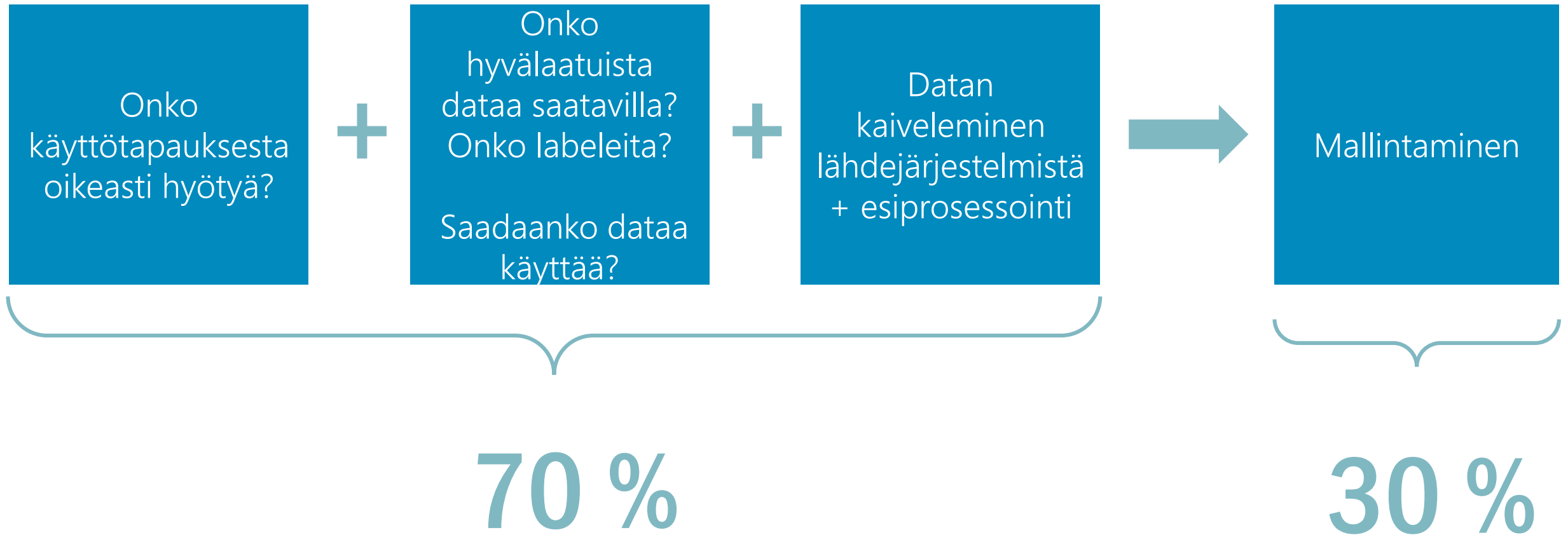


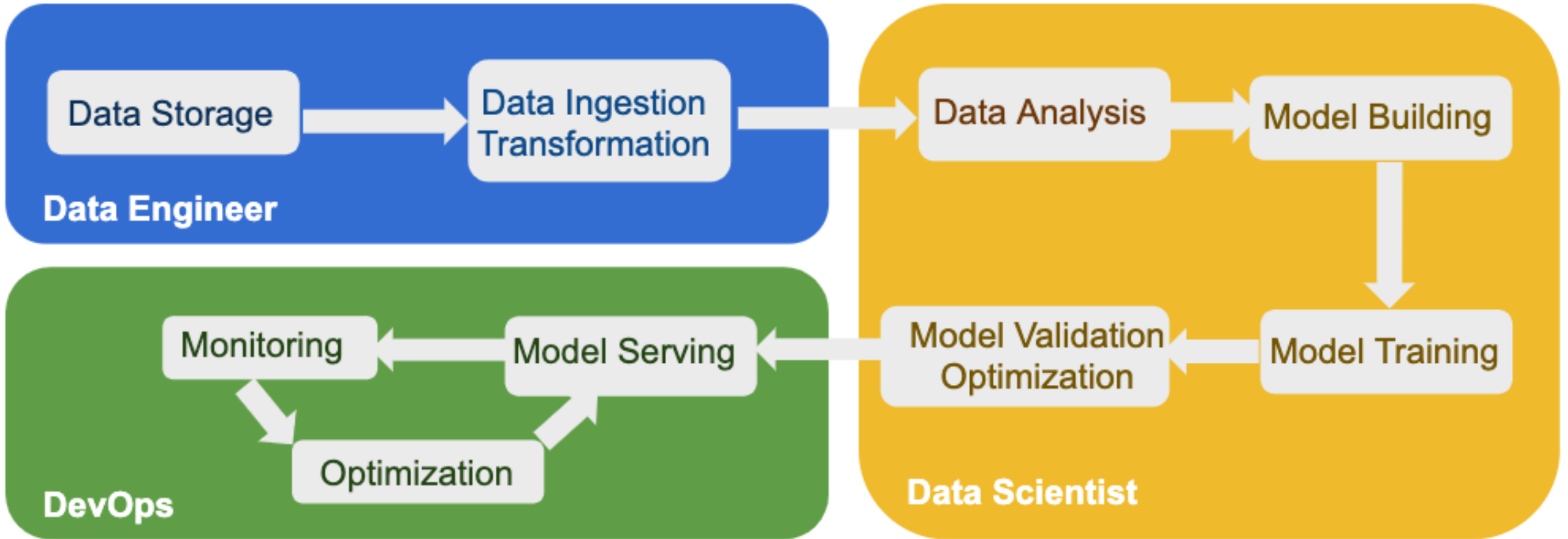
Tekoälymallin elinkaari ideasta tuotantoon #MLOps

Kela|Fpa 
KelaLAB

Tekoälymallin koulutus in a nutshell

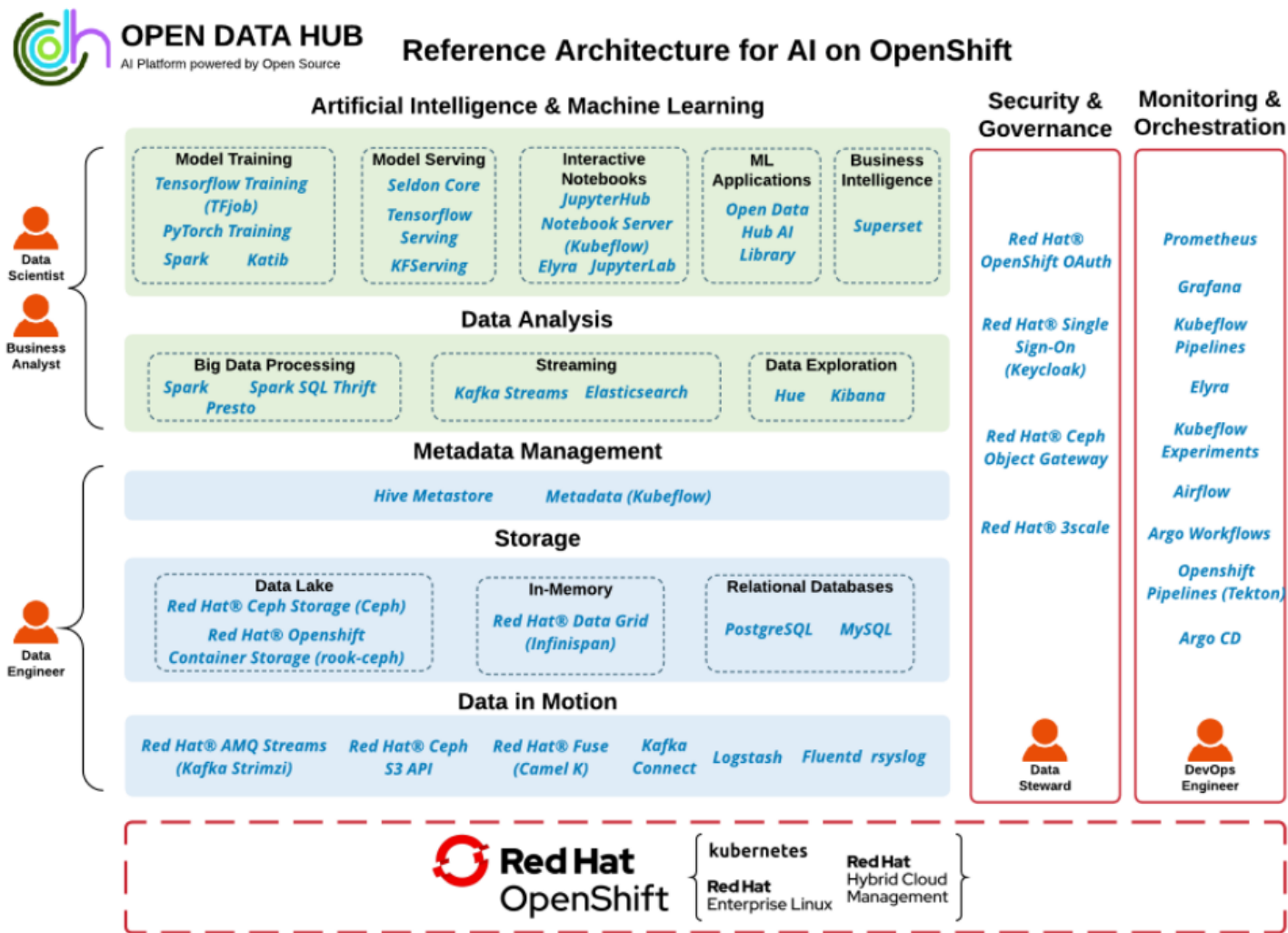


Tekoälymallin tuotannollistaminen



[Architecture · OpenDataHub](#)

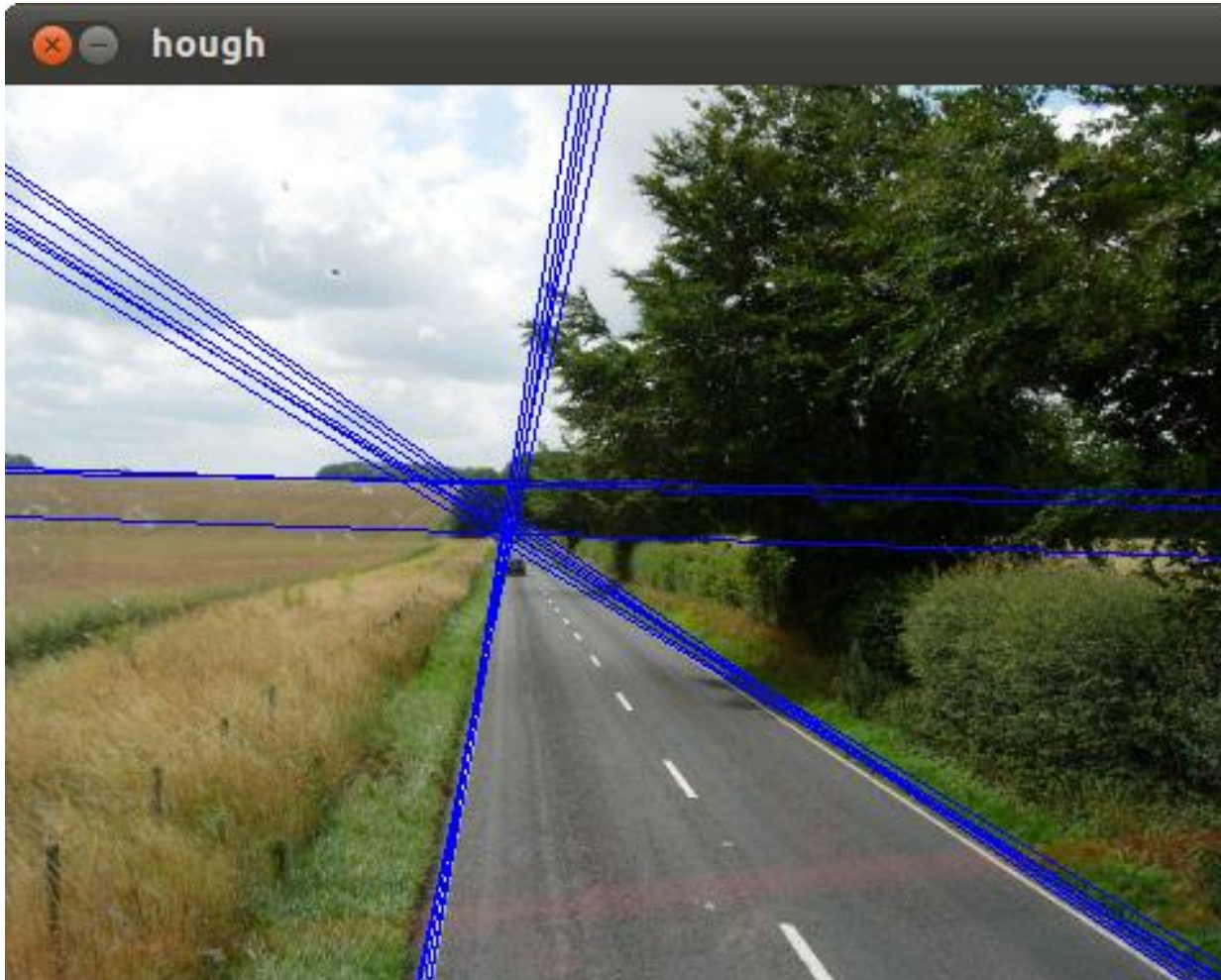
Open Data Hub – opensource-data science –työkalupakki OpenShiftille



Käyttötapauksia matkan varrelta

Kela|Fpa[®]
KelaLAB

Kuvien kääntäminen



- Tunnistetaan dokumentin viivojen "pääsuunta" Hough-muunnoksella
- Tunnistetaan Hough-muunnoksen jälkeen, onko kuva 0, 90, 180 vai 270 asteen kulmassa.
 - Tunnistus voidaan tehdä hyödyntäen OCR:ää tai esim. konvoluutioneuroverkolla

Chättibotit

- <https://www.kela.fi/chattirobotti>
- Taustalla Boost.ai:n botti, jota hyödynnetään meillä on-premise –ympäristössä
 - Boostin käyttämä kielimalli lienee multilingual BERT, mutta varmaa tietoa ei ole.
 - Kouluttaminen omalla datalla onnistuu (Boostin malli tehty Pytorchilla)

Älykäs OCR (DocCVR)

The image displays two versions of the Form W-4, "Employee's Withholding Certificate".

Left Example (Scanned):

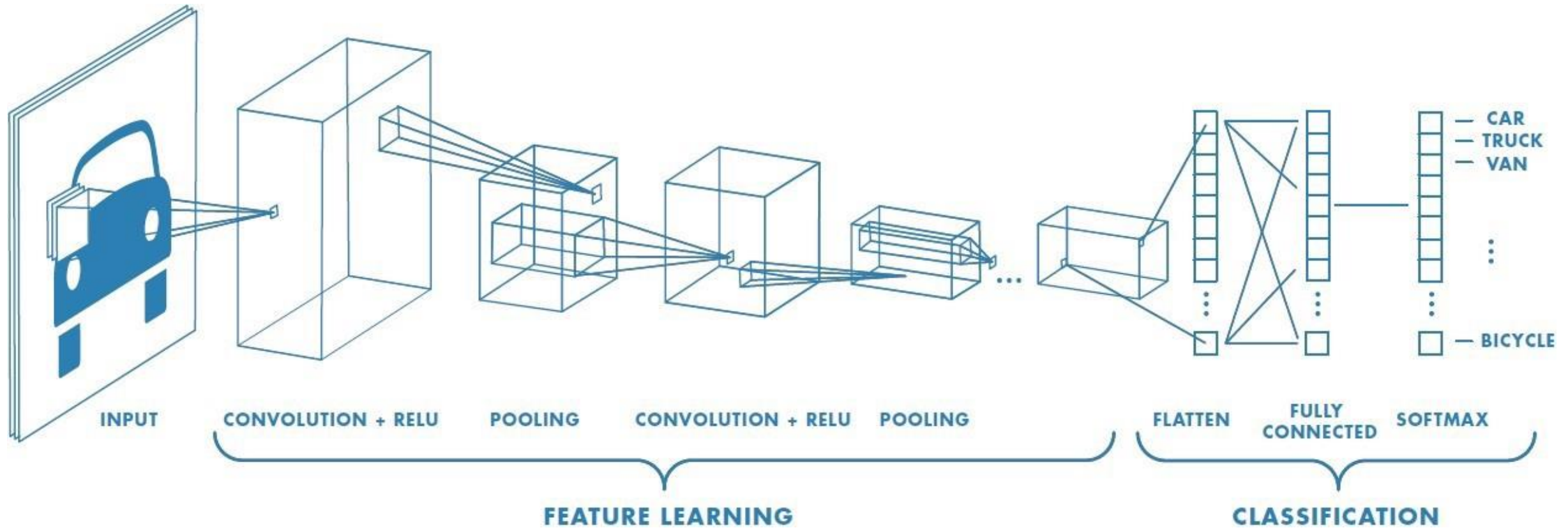
- Form W-4**
- Step 1: Enter Personal Information**
 - (a) First name and middle initial: Actrian
 - (b) Last name: Rosebrock
 - (c) Social security number: 12-34-56789
 - (d) Date of birth: 06/10/2020
 - (e) City or town, state, and ZIP code: Philadelphia, PA 19109
- Step 2: Multiple Jobs or Spouse Works**
 - (a) Complete this step if you (1) hold more than one job at a time, or (2) are married filing jointly and your spouse also works. The correct amount of withholding depends on income earned from all of these jobs.
 - (b) Only one of the following:
 - (i) Use the estimator at www.irs.gov/W4App for most accurate withholding for this step (and Steps 3-4), or
 - (ii) Use the Multiple Jobs Worksheet on page 3 and enter the result in Step 4(b) below for roughly accurate withholding, or
 - (iii) If there are only two jobs total, you may check this box. Do the same on Form W-4 for the other job. This option is accurate for jobs with similar pay; otherwise, more tax than necessary may be withheld.
- Step 3: Claim Dependents**
 - (a) If your income will be \$200,000 or less (\$400,000 or less if married filing jointly):
 - Multiply the number of qualifying children under age 17 by \$2,000: 3 x \$2,000 = \$6,000
 - Multiply the number of other dependents by \$500: 1 x \$500 = \$500
 - Add the amounts above and enter the total here: \$6,500
- Step 4 (optional): Other Adjustments**
 - (a) Other income (not from jobs). If you want tax withheld for other income you expect this year that won't have withholding, enter the amount of other income here. This may include interest, dividends, and retirement income: \$400
 - (b) Deductions. If you expect to claim deductions other than the standard deduction and want to reduce your withholding, use the Deductions Worksheet on page 3 and enter the result here: \$400
 - (c) Extra withholding. Enter any additional tax you want withheld each pay period: \$400
- Step 5: Sign Here**
 - Under penalties of perjury, I declare that this certificate, to the best of my knowledge and belief, is true, correct, and complete.
 - Employee's signature (This form is not valid unless you sign it.): Actrian Rosebrock
 - Date: 2020/06/10
- Employers Only**
 - Employer's name and address: Pymagelarch, PO BOX 1234, Philadelphia, PA 19109
 - First date of employment: 12-34-56789
 - Employer identification number (EIN): 12-3456789

Right Example (Digital Form):

- Form W-4**
- Employee's Withholding Certificate**
- Step 1: Enter Personal Information**
 - (a) First name and middle initial: [Blank]
 - (b) Last name: [Blank]
 - (c) Social security number: [Blank]
 - (d) Date of birth: [Blank]
 - (e) City or town, state, and ZIP code: [Blank]
- Step 2: Multiple Jobs or Spouse Works**
 - (a) Complete this step if you (1) hold more than one job at a time, or (2) are married filing jointly and your spouse also works. The correct amount of withholding depends on income earned from all of these jobs.
 - (b) Only one of the following:
 - (i) Use the estimator at www.irs.gov/W4App for most accurate withholding for this step (and Steps 3-4), or
 - (ii) Use the Multiple Jobs Worksheet on page 3 and enter the result in Step 4(b) below for roughly accurate withholding, or
 - (iii) If there are only two jobs total, you may check this box. Do the same on Form W-4 for the other job. This option is accurate for jobs with similar pay; otherwise, more tax than necessary may be withheld.
- Step 3: Claim Dependents**
 - (a) If your income will be \$200,000 or less (\$400,000 or less if married filing jointly):
 - Multiply the number of qualifying children under age 17 by \$2,000: 3 x \$2,000 = \$6,000
 - Multiply the number of other dependents by \$500: 1 x \$500 = \$500
 - Add the amounts above and enter the total here: \$6,500
- Step 4 (optional): Other Adjustments**
 - (a) Other income (not from jobs). If you want tax withheld for other income you expect this year that won't have withholding, enter the amount of other income here. This may include interest, dividends, and retirement income: \$400
 - (b) Deductions. If you expect to claim deductions other than the standard deduction and want to reduce your withholding, use the Deductions Worksheet on page 3 and enter the result here: \$400
 - (c) Extra withholding. Enter any additional tax you want withheld each pay period: \$400
- Step 5: Sign Here**
 - Under penalties of perjury, I declare that this certificate, to the best of my knowledge and belief, is true, correct, and complete.
 - Employee's signature (This form is not valid unless you sign it.): [Blank]
 - Date: [Blank]
- Employers Only**
 - Employer's name and address: [Blank]
 - First date of employment: [Blank]
 - Employer identification number (EIN): [Blank]

- Tunnistetaan dokumenttityyppi
- Tunnistetaan dokumenttipohjan versio
- Etsitään keypointit (ORB, SIFT, AKAZE)
- Mätsätään keypointit (RANSAC)
- Lasketaan homografiamatriisi ja korjataan kuvan asento

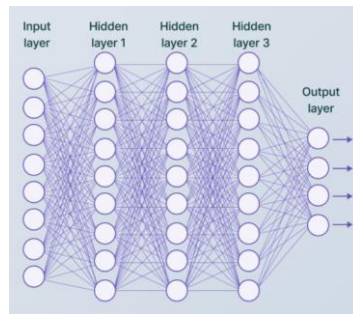
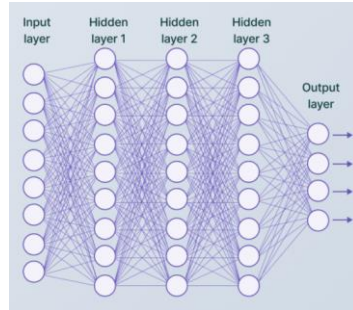
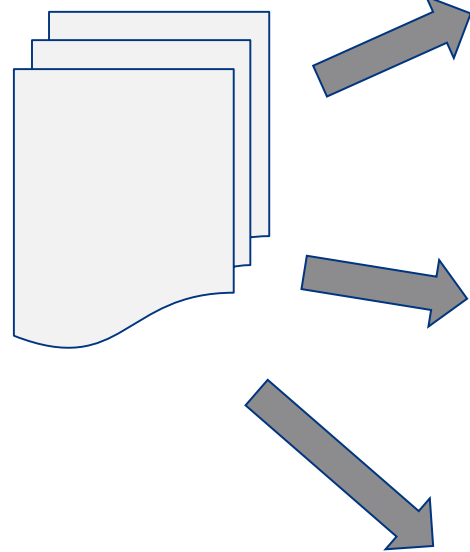
Dokumenttien automaattinen tunnistaminen (1/4)



Dokumenttien automaattinen tunnistaminen (2/4)

Convolutional neural networks

Documents to be classified



Text classification models

Random forest,
logistic regression
etc.

Ensembling

Classification

"Smart" OCR

Dokumenttien automaattinen tunnistaminen (3/4)

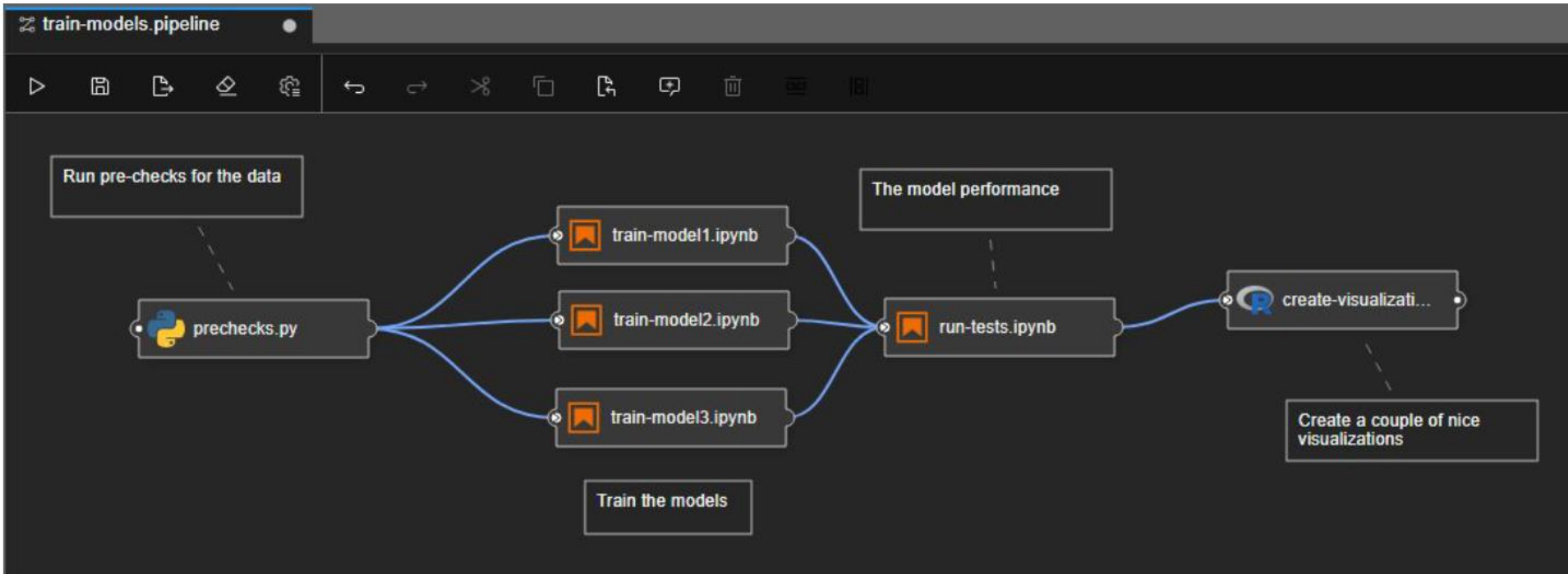
```
1  apiVersion: machinelearning.seldon.io/v1
2  kind: SeldonDeployment
3  metadata:
4    name: {{ .Values.sd.metadata.name }}
5    namespace: {{ .Values.common.namespace }}
6  spec:
7    name: {{ .Values.sd.spec.name }}
8    oauth_key: oauth-key
9    oauth_secret: oauth-secret
10   predictors:
11     - componentSpecs:
12       - spec:
13         containers:
14           ...|
15
16         initContainers:
17           - args:
18             - '-c'
19             - cp -r /models/* /mnt/models/predtype/; ls -R /mnt/models/;
20             ...
21
22         volumes:
23           ...
24
25   graph:
26     name: preprocess
27     type: TRANSFORMER
28     endpoint:
29       type: {{ .Values.sd.graph.endpoint.type }}
30     children:
31       - name: combiner
32         type: COMBINER
33         children:
34           - name: cnn
35             type: MODEL
36             endpoint:
37               type: {{ .Values.sd.graph.endpoint.type }}
38             children: []
39           - name: text
40             type: MODEL
41             endpoint:
42               type: {{ .Values.sd.graph.endpoint.type }}
43             children: []
44   svcOrchSpec:
45     ...
46
47   name: {{ .Values.sd.graph.name }}
```

Esimerkki ensemble-
koneoppimismallin
julkaisemisesta käyttäen
Seldon Corea (yaml-filua
on lyhennetty kohdissa,
joissa ...)

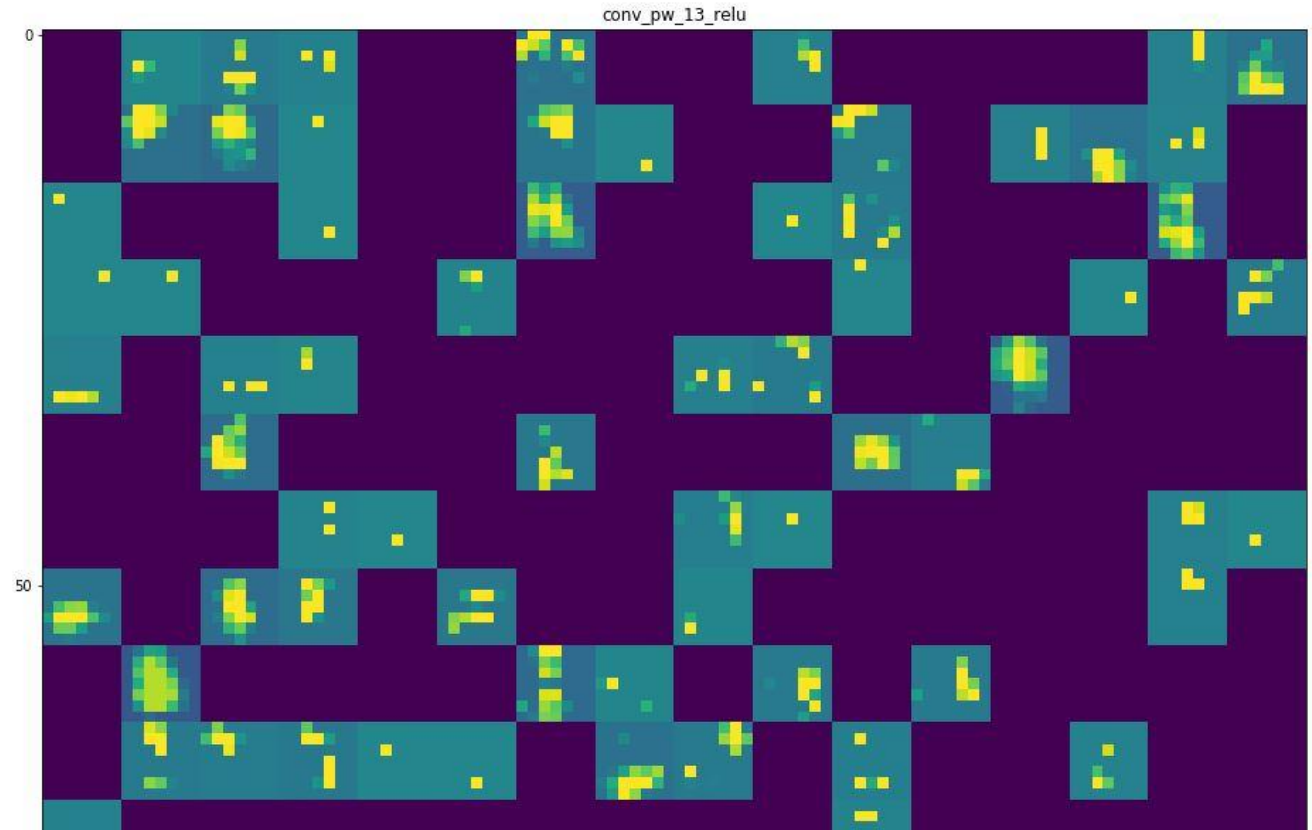
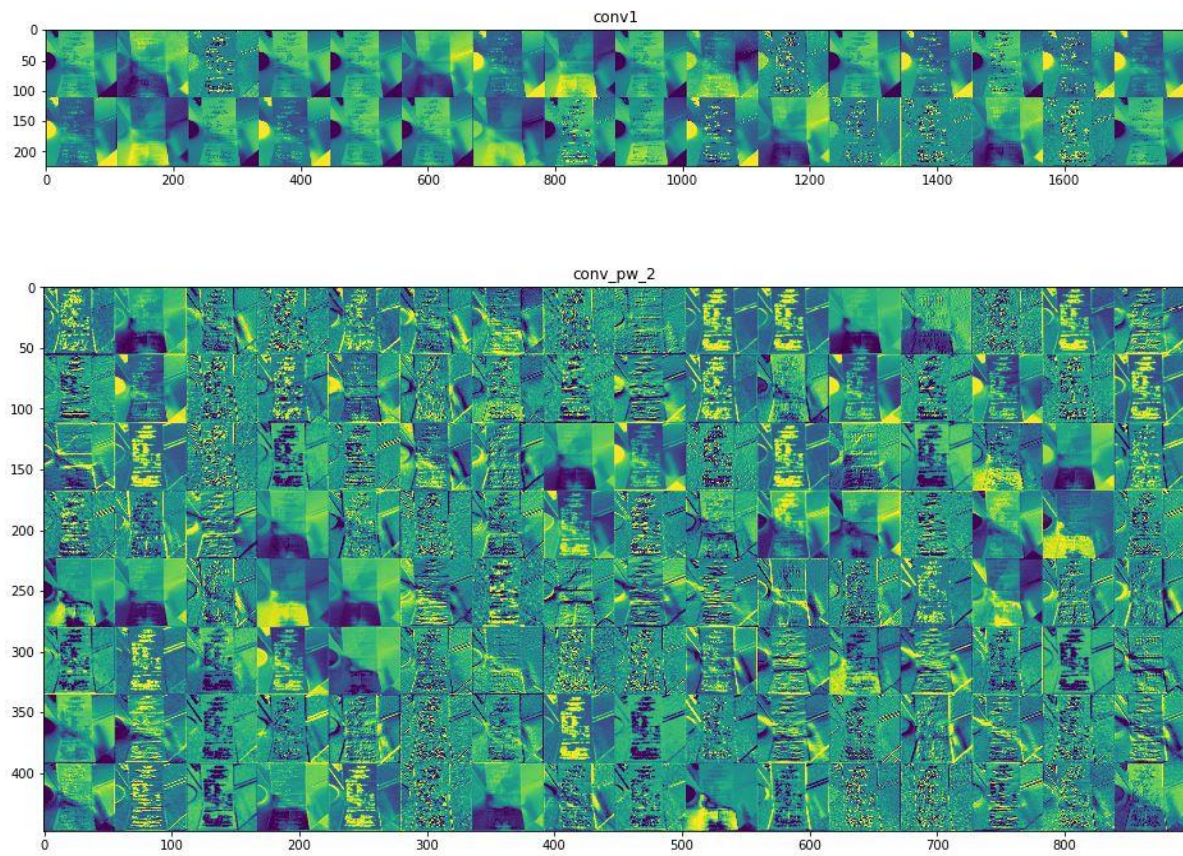
Normaali Helm-chart
ArgoCD:llä

Huomaa inference-
graafin eri komponentit

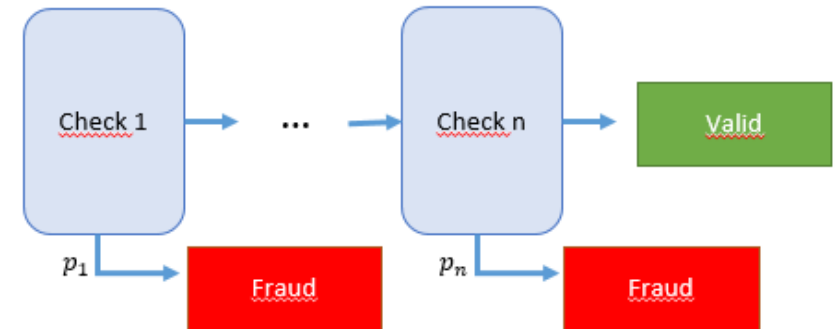
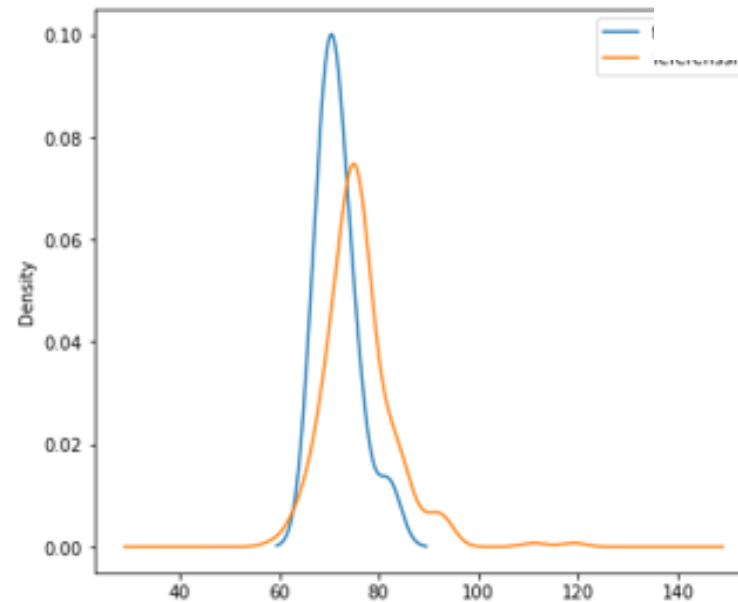
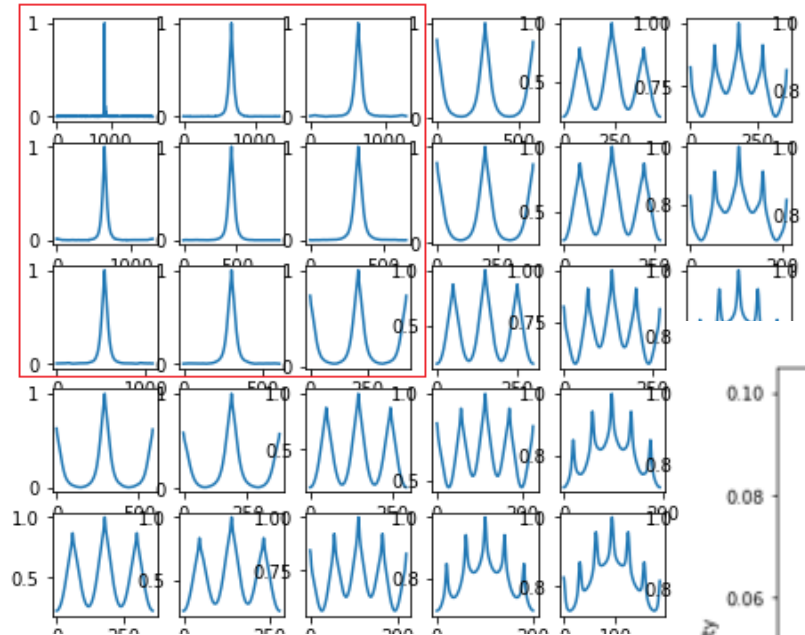
Dokumenttien automaattinen tunnistaminen (4/4)



Konvoluutoneuroverkot (4/4)



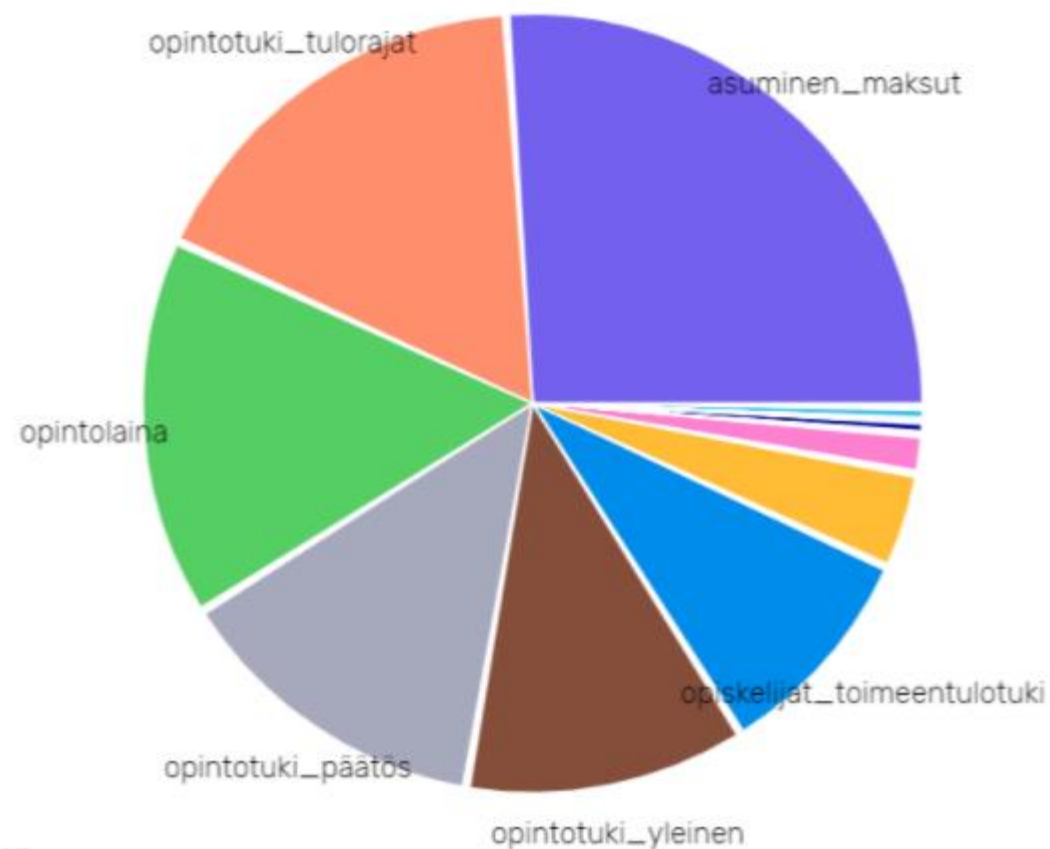
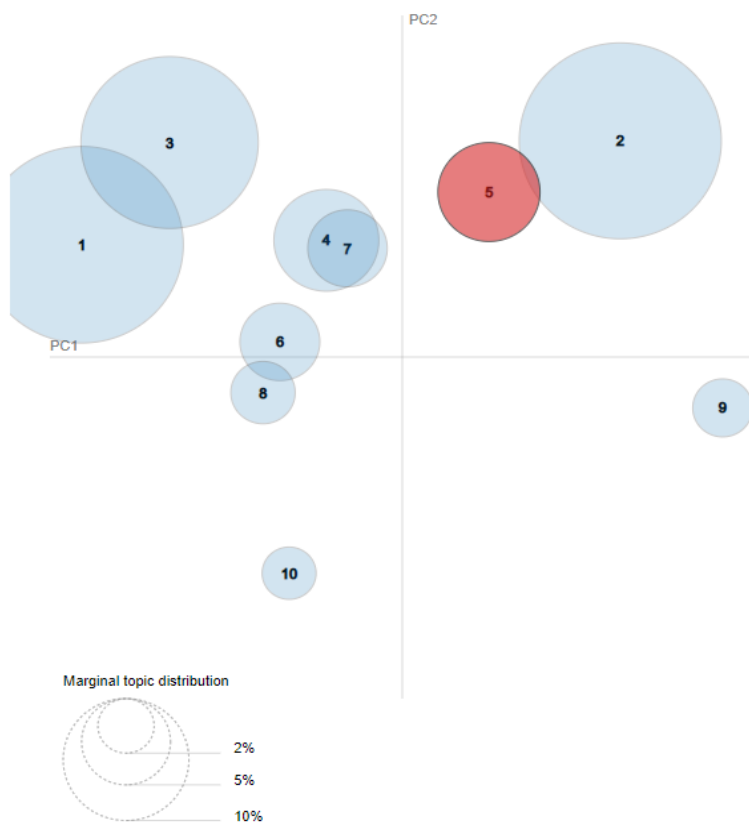
Väärennettyjen asiakirjojen tunnistaminen



Asiakaspuhuteluiden analysointi

Selected Topic:

Intertopic Distance Map (via multidimensional scaling)



Kiitos! Kysymyksiä,
kommentteja?

IT-innovaatioyksikkö

Kela|Fpa 