

# AI Model til Forudsigelse af Akutte Psykiatriske Genindlæggelser: Fra forskningsmiljø til Implementering i Sundhedsplatformen

Anders C. Riis-Jensen, Ph.D. teoretisk fysik

Teamleder og dataanalytiker

Team Prædiktiv Modellering

Region Hovedstaden

Kontakt: [anders.christian.riis-jensen.01@regionh.dk](mailto:anders.christian.riis-jensen.01@regionh.dk)

# Sundhedsplatformen (SP)

## SP og Østdanmark:

- Ca. 50.000 slutbrugere
- Ca. 3 mio. patienter
- Ca. 110.000 arbejdsstationer & mobile enheder
- Ca. 3.500 medicinsk udstyr
- CPR nummeret er unikt

## Data (Februar 2021):

- ❖ SP databasen indeholder 320.581.128.052 rækker
- ❖ fordelt på 11.774 tabeller



# Fra forskningsidé til prædiktiv model i Sundhedsplatformen

**Forskningsidé**

- Definition af projekt
- Definition af succes

**Møde**

- Forventning
- Jura
- Finansiering
- Tekniske krav
- Data tilgængelighed
- Tidsplan for implementering
- Ansvar

**Definitioner:**

- Tidsestimering og finansiering er angivet fra 1-5 i kasserne:  
1 2 3 4 5
- Basis-finansiering af SP Sundhedsdata: ■
- Minimum ekstern finansiering: ■
- Muligt eller nødvendigt tilkøb: ■

### EPIC model Implementering

- Dialog med EPIC**
  - Dækker EPIC model behov?
  - Model performance i dansk kontekst
- Target definition**
  - Definition af Target for prædiktion
- Mapping af datainput**
  - Dialog med klinikere for at sikre at model input matcher dansk kontekst
- Teknisk opsætning og Validering**
  - CE mærkning
  - SP opsætning
  - Klinisk validering
- Visning i Hyperspace**
  - Rapport
  - Patientliste
  - Patientjournal
  - Andet sted
- Monitorering**
  - Kontinuerlig monitorering af model performance

### SP-Sundhedsdata modeludvikling

- Møde med klinikere**
  - Definition af Target for prædiktion
  - Definition af succes
  - Vigtige features for prædiktion
- Data ekstrakt**
  - Udvælgelse af data fra SP og upload til analyseplatform
- Modeludvikling**
  - Feature selektion
  - Udvikling af AI model af SP Sundhedsdata

### Forsker modeludvikling

- Møde med data scientists**
  - Intro til analyseplatform
- Data ekstrakt**
  - Udvælgelse af data fra SP og upload til analyseplatform
- Modeludvikling**
  - Forskere udvikler model
  - Support kan tilkøbes

### Implementering

- Teknisk setup**
  - SP Databasebyg
  - Modelimplementering
- Klinisk validering**
  - Godkendelse af klinikere
  - Juridisk godkendelse og dokumentation
- Visning i Hyperspace**
  - Rapport
  - Patientjournal
  - Patientliste
  - Andet sted
- Monitorering**
  - Kontinuerlig monitorering af model performance

# Fra forskningsidé til prædiktiv model i Sundhedsplatformen

**Forskningsidé**

- Definition af projekt
- Definition af succes

**Møde**

- Forventning
- Jura
- Finansiering
- Tekniske krav
- Data tilgængelighed
- Tidsplan for implementering
- Ansvar

**Definitioner:**

- Tidsestimering og finansiering er angivet fra 1-5 i kasserne:  
1 2 3 4 5
- Basis-finansiering af SP Sundhedsdata: ■
- Minimum ekstern finansiering: ■
- Muligt eller nødvendigt tilkøb: ■

### EPIC model Implementering

- Dialog med EPIC**
  - Dækker EPIC model behov?
  - Model performance i dansk kontekst
- Target definition**
  - Definition af Target for prædiktion
- Mapping af datainput**
  - Dialog med klinikere for at sikre at model input matcher dansk kontekst
- Teknisk opsætning og Validering**
  - CE mærkning
  - SP opsætning
  - Klinisk validering
- Visning i Hyperspace**
  - Rapport
  - Patientliste
  - Patientjournal
  - Andet sted
- Monitorering**
  - Kontinuerlig monitorering af model performance

### SP-Sundhedsdata modeludvikling

- Møde med klinikere**
  - Definition af Target for prædiktion
  - Definition af succes
  - Vigtige features for prædiktion
- Data ekstrakt**
  - Udvælgelse af data fra SP og upload til analyseplatform
- Modeludvikling**
  - Feature selektion
  - Udvikling af AI model af SP Sundhedsdata

### Forsker modeludvikling

- Møde med data scientists**
  - Intro til analyseplatform
- Data ekstrakt**
  - Udvælgelse af data fra SP og upload til analyseplatform
- Modeludvikling**
  - Forskere udvikler model
  - Support kan tilkøbes

### Implementering

- Teknisk setup**
  - SP Databasebyg
  - Modelimplementering
- Klinisk validering**
  - Godkendelse af klinikere
  - Juridisk godkendelse og dokumentation
- Visning i Hyperspace**
  - Rapport
  - Patientjournal
  - Patientliste
  - Andet sted
- Monitorering**
  - Kontinuerlig monitorering af model performance

# AI modeller i Sundhedsplatformen

**Risiko for patientudeblivelse** [*Gyldenkerne et. al. 2020*]

- Prædikterer risikoen for patientudeblivelse
- Baseret på 10 input variable inklusiv, højde, vægt, køn, postnummer, lead time, aftalelængde...
- Retunerer et enkel tal som mål for risiko for udeblivelse
- Simpelt kunstigt neuralt netværk udviklet (Keras, Python)
- **Videreudvikles lige nu til skaleret version til brug på alle afdelinger på tværs af de to regioner**

# Risiko for udeblivelse

Output i Sundhedsplatformen:

Tidspunkt	Afsnit	CPR	Navn	Kontakt nummer	Besøgstype	Ressource	Tolk	Transport	Risiko for udeblivelse
15-10-2021 09:00	M-Y	XXXXXX-XXXX	Tante Ursula	12345678	Konsultation	Læge			44 %
15-10-2021 09:00	M-Y	XXXXXX-XXXX	Benny	56781234	Lunge- funktionstest	Sygeplejerske			87 %
15-10-2021 09:15	M-X	XXXXXX-XXXX	John	87654321	Blodprøve	Sygeplejerske	Fransk		20 %
...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...

# AI modeller i Sundhedsplatformen

## Risiko for patientudeblivelse *[Gyldenkærne et al. 2020]*

- Prædikterer risikoen for patientudeblivelse
- Baseret på 10 input variable inklusiv, højde, vægt, køn, postnummer, lead time, aftalelængde...
- Retunerer et enkel tal som mål for risiko for udeblivelse
- Simpelt kunstigt neuralt netværk udviklet (Keras, Python)
- **Videreudvikles lige nu til skaleret version til brug på alle afdelinger på tværs af de to regioner**

## Infektioner I patienter med kronisk lymfatisk leukæmi (CLL) *[CLL-TIM, Agius et al., Nat Comm, 2020]*

- Prædikterer risikoen for fremtidig livstruende infektion I patienter diagnosticeret med kronisk lymfatisk leukæmi (CLL)
- Pilot-implementering på Rigshospitalet
- Ensemble model bestående af 28 AI algoritmer (python)
- Bl.a. baseret på laboratoriesvar, familiær historie, cancer staging
- Retunerer risiko og konfidens sammen med en række risikofaktorer

# Kronisk lymfatisk leukæmi

Output i Sundhedsplatformen:

Name and CPR	Anna – XXXXXX-XXXX
Risk	High
Confidence	High
Risk factor 1 – high score	Binet Stage B,C or unknown
Risk factor 2 – high score	Platelets - Value
Risk factor 3 – high score	Absolute Erythrocyte Count
Risk factor 1 – low score	Age
Risk factor 2 – low score	Creatinine - value
Risk factor 3 – low score	IGHV unmutated



## Projekter under implementering i SP

## Projekter under implementering i SP

### Early detection of sepsis – EPIC model

- Tidlig advarsel (6 timer) omkring udvikling af sepsis
- Baseret på laboratoriesvar, puls, temperature, respiration mm.
- Bliver i øjeblikket performance testet på SPs patientdatabase
- Videre randomiseret studie følger hvis performance er høj nok

## Projekter under implementering i SP

### Early detection of sepsis – EPIC model

- Tidlig advarsel (6 timer) omkring udvikling af sepsis
- Baseret på laboratoriesvar, puls, temperature, respiration mm.
- Bliver i øjeblikket performance testet på SPs patientdatabase
- Videre randomiseret studie følger hvis performance er høj nok

### Iskæmiske hjertesmerter

- Prædiktiv overlevelseskurve over kommede måneder og år
- Bl.a. baseret på laboratoriesvar, familiær historie, EKG, puls, temperatur
- Bliver også implementeret i Norge og Island

## Projekter under implementering i SP

### Akutte psykiatriske genindlæggelser

- Prædikterer risikioen for en akut genindlæggelse for psykiatrisk indlagte patienter
- I øjeblikket under udvikling i samarbejde med forskningsgruppe fra PC København
- Planlagt første klinisk implemetering slut 2022

## Hvad er udfordringen?

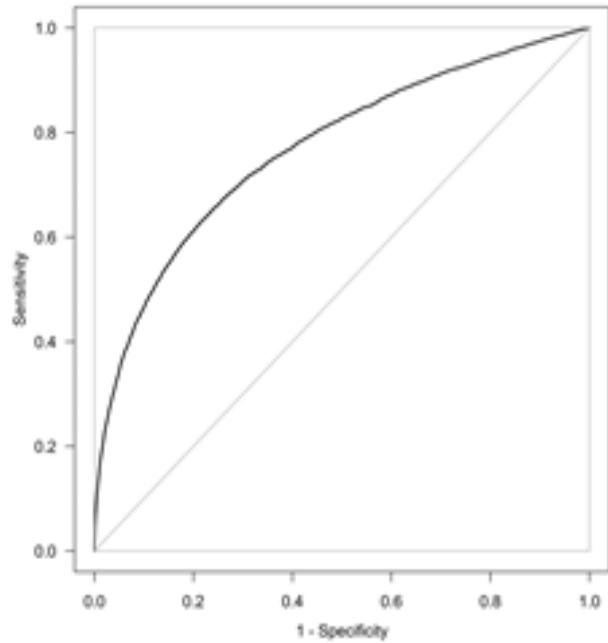
- Massere af forskningsmæssig udvikling som ikke bliver anvendt I praksis
  - *Modeloutput skal vises inde I EPJ systemet for optimal brug*

## Prædiktion af akutte psykiatriske genindlæggelser (initialt arbejde på registre data)

**Baggrund:** akutte genindlæggelser der ikke er planlagte anses for uhensigtsmæssige og der er et ønske om at reducere forekomsten for at forbedre behandlingen

**Formål:** Identificere personer med høj risiko for akutte genindlæggelse inden for 30 dage efter udskrivning fra en psykiatrisk afdeling

# Prædiktion af akutte psykiatriske genindlæggelser (initialt arbejde på registre data)



## Acute Psychiatric Re-admission Risk at Discharge (APRAD)



Diagnosis	Number of Acute Psychiatric Re-admissions last 12 months							
	0		1		2-5		>5	
Disorder of personality and behavior (F60-69)	16	14	24	28	41	52	71	79
Schizophrenia and related disorders (F20-29)	14	20	24	34	40	47	68	77
Affective disorders (F30-39)	11	16	16	38	26	53	64	84
Disorder due to substance use (F10-19)	16	14	25	28	32	46	63	66
Other mental disorder diagnosis	13	13	22	26	32	49	60	70
	<4	1-4	<4	1-4	<4	1-4	<4	1-4

Absolute risk of acute psychiatric re-admission	Length of hospital stay (days)					
10-14%	15-24%	25-34%	35-44%	45-64%	≥65%	

Clinical Risk chart	1-factor*		3-factor**		5-factor***	
	ROC-AUC	95%CI	ROC-AUC	95%CI	ROC-AUC	95%CI
2015 (training data)	0.728	0.722-0.734	0.745	0.738-0.751	0.747	0.741-0.754
1995-2015	NA	NA	0.724	0.722-0.725	0.730	0.728-0.731
2016 (test data)	<b>0.733</b>	0.727-0.739	<b>0.749</b>	0.743-0.755	<b>0.751</b>	0.745-0.756

## Steps i udvikling og implementering af AI model

- **Forskningsidé**



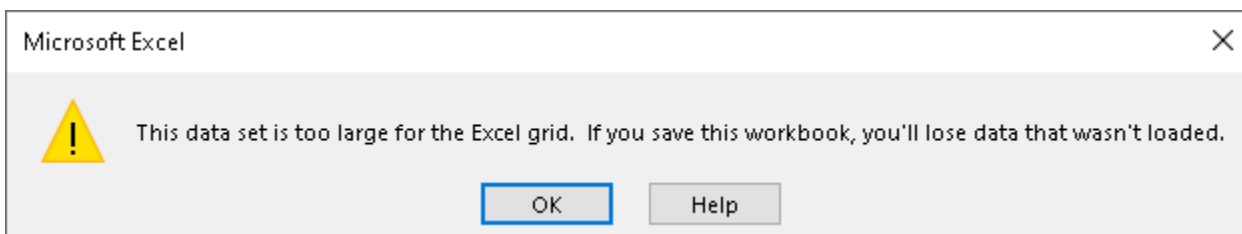


## Steps i udvikling og implementering af AI model

- **Forskningsidé**



- **Data**



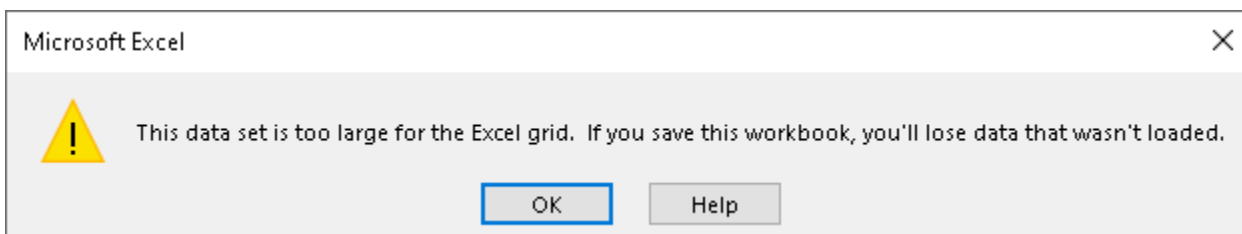
## Steps i udvikling og implementering af AI model

- **Forskningsidé**



- **Data**

- **Dataanalyse og modeludvikling**



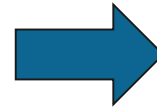
# Steps i udvikling og implementering af AI model

- Forskningsidé

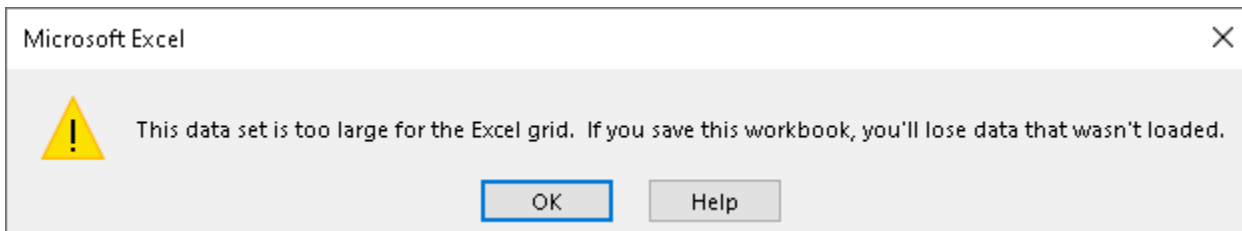


- Data

- Dataanalyse og modeludvikling



- Backend setup i SP



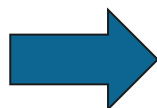
## Steps i udvikling og implementering af AI model

- **(Gen)validering på SPs database**



## Steps i udvikling og implementering af AI model

- (Gen)validering på SPs database

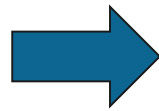


- Byg af visning i SP

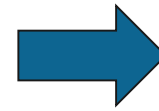


## Steps i udvikling og implementering af AI model

- (Gen)validering på SPs database



- Byg af visning i SP



- Monitorering af modelperformance



*"It pains me to tell you this, but it ain't broke."*

## Steps i udvikling og implementering af AI model

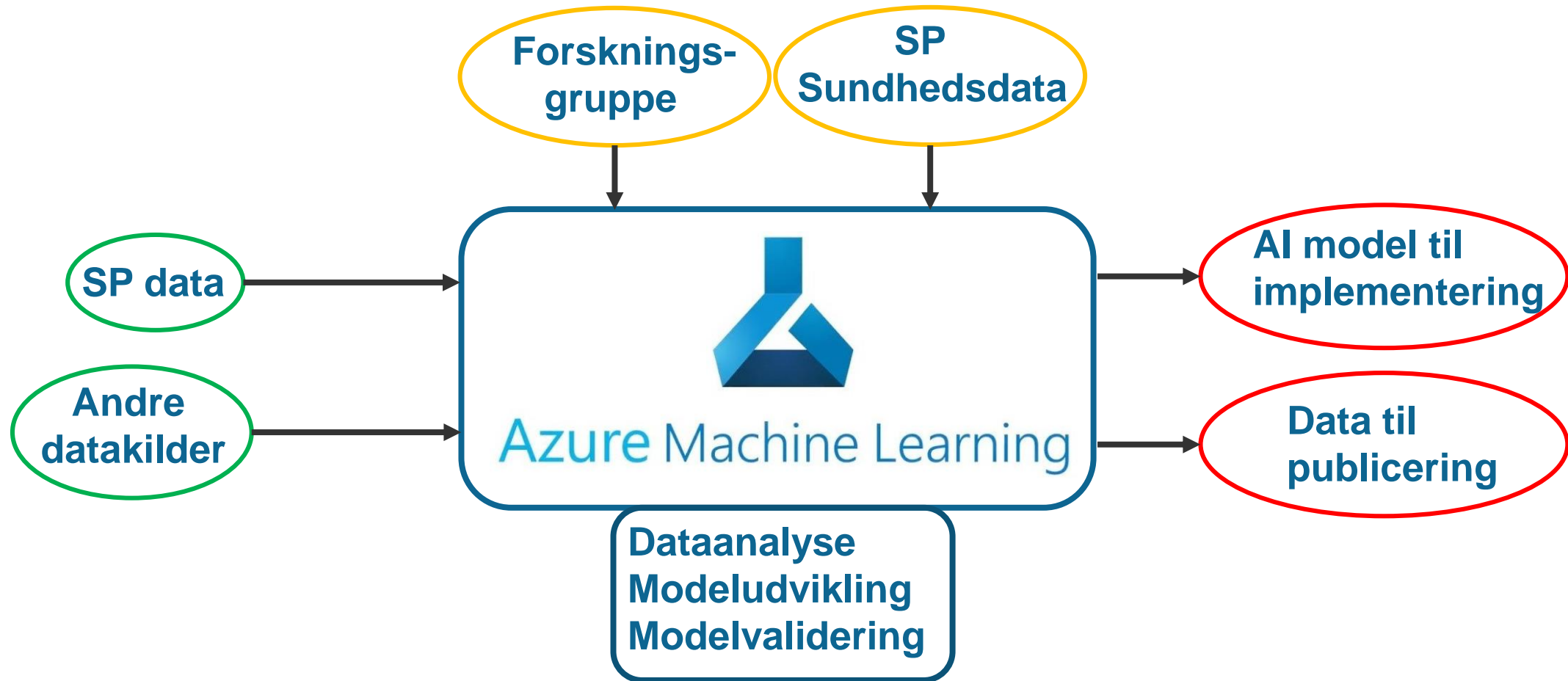
- **Forskningsidé**
- **Data**
- **Dataanalyse og modeludvikling**
- **Backend setup i SP**
- **(Gen)validering på SPs database**
- **Byg af visning i SP**
- **Monitorering af modelperformance**

## Steps i udvikling og implementering af AI model

- Forskningsidé
- Data
- Dataanalyse og modeludvikling
- Backend setup i SP
- (Gen)validering på SPs database
- Byg af visning i SP
- Monitorering af modelperformance

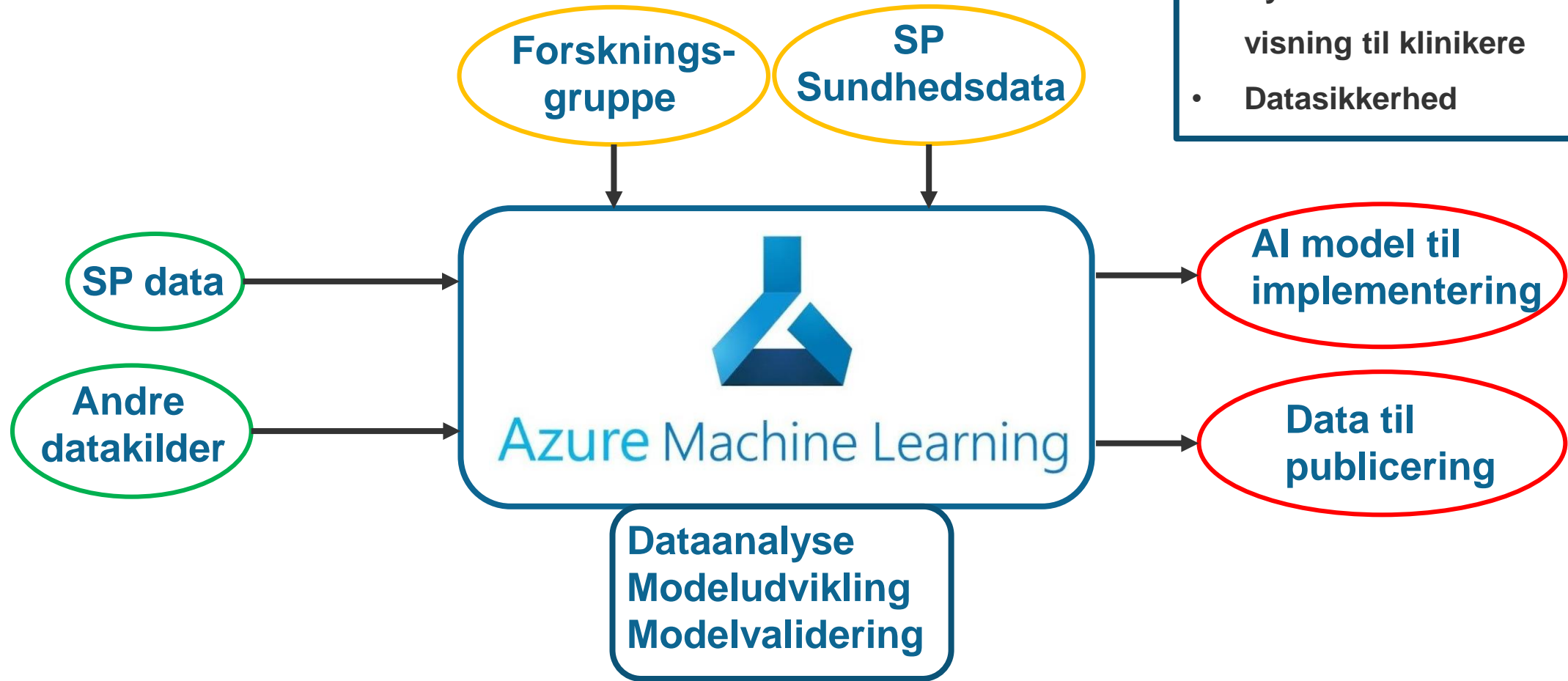


# Roadmap til implementering af AI modeller: fra forskningsmiljø til Sundhedsplatformen



# Roadmap til implementering af AI modeller: fra forskningsmiljø til Sundhedsplatformen

- Øget datakvalitet
- Hurtigere implementering i SP
- Symbiose mellem AI model og visning til klinikere
- Datasikkerhed





Region Hovedstaden



# Welcome to the Azure Machine Learning Studio

[Create workspace](#)

Create a new workspace, or open one of your recent workspaces to pick up where you left off.

## Recent workspaces

forskerpl-7mk0mr-mlw

Subscription  
COK-Prod-Forskerplatform

forskerpl-p2ss0q-mlw

Subscription  
COK-Prod-Forskerplatform

spdatateuwfpml

Subscription  
COK-Test-Forskerplatform

[View all workspaces →](#)

## Learning components



Build AI solutions with Azure  
Machine Learning



Introduction to the Azure  
Machine Learning SDK



Train a machine learning model  
with Azure ML

[View all learning components →](#)

## Tutorials



What is Azure Machine  
Learning?



Train your first ML model with  
Notebook



Create, explore and deploy  
Automated ML experiments.

[View all tutorials →](#)



Region Hovedstaden

+ New

Home

Author

Notebooks

Automated ML

Designer

Assets

Data

Jobs

Components

Pipelines

Environments

Models

Endpoints

Manage

Compute

Datstores

Linked Services

Data Labeling

Region Hovedstaden &gt; forskerpl-p2ss0q-mlw

## forskerpl-p2ss0q-mlw



Create new ▾



## Notebooks

Code with Python SDK and run sample experiments.

Start now



## Automated ML

Automatically train and tune a model using a target metric.

Start now



## Designer

Drag-and-drop interface from prepping data to deploying models.

Start now

## Recent resources

Jobs

Compute

Models

Data

Display name



Experiment

Status

Logs

Submitted time

Submitte...

Job type



No jobs to display

Region Hovedstaden > forskerpl-p2ss0q-mlw > C

## Compute

Compute instances **Compute clusters** Inference

+ New Refresh Start Stop Re

Search

Name	☆	State
medium		Stop
small		Stop
Datacheck		Stop

## Create compute instance

 Required Settings Advanced Settings

<input type="radio"/> Standard_D11 2 cores, 14GB RAM, 100GB storage	Memory optimized	58 cores	\$0.22/hr
<input type="radio"/> Standard_D11_v2 2 cores, 14GB RAM, 100GB storage	Memory optimized	58 cores	\$0.19/hr
<input type="radio"/> Standard_D12 4 cores, 28GB RAM, 200GB storage	Memory optimized	58 cores	\$0.44/hr
<input type="radio"/> Standard_D12_v2 4 cores, 28GB RAM, 200GB storage	Memory optimized	58 cores	\$0.38/hr
<input type="radio"/> Standard_D13 8 cores, 56GB RAM, 400GB storage	Memory optimized	58 cores	\$0.88/hr
<input type="radio"/> Standard_D13_v2 8 cores, 56GB RAM, 400GB storage	Memory optimized	58 cores	\$0.76/hr
<input type="radio"/> Standard_D14 16 cores, 112GB RAM, 800GB storage	Memory optimized	58 cores	\$1.77/hr
<input type="radio"/> Standard_D14_v2 16 cores, 112GB RAM, 800GB storage	Memory optimized	58 cores	\$1.52/hr
<input type="radio"/> Standard_D15_v2 20 cores, 140GB RAM, 1000GB storage	Memory optimized	58 cores	\$1.90/hr
<input type="radio"/> Standard_D16_v3 16 cores, 64GB RAM, 400GB storage	General purpose	58 cores	\$0.96/hr
<input type="radio"/> Standard_D16ds_v4 16 cores, 64GB RAM, 600GB storage	General purpose	58 cores	\$1.09/hr



Region Hovedstaden

+ New

Home

Author

Notebooks

Automated ML

Designer

Assets

Data

Jobs

Components

Pipelines

Environments

Models

Endpoints

Manage

Compute

Datstores

Linked Services

Data Labeling

Region Hovedstaden &gt; forskerpl-p2ss0q-mlw

## forskerpl-p2ss0q-mlw



Create new ▾



## Notebooks

Code with Python SDK and run sample experiments.

Start now



## Automated ML

Automatically train and tune a model using a target metric.

Start now



## Designer

Drag-and-drop interface from prepping data to deploying models.

Start now

## Recent resources

Jobs

Compute

Models

Data

Display name



Experiment

Status

Logs

Submitted time

Submitte...






Job type



No jobs to display

Region Hovedstaden &gt; forskerpl-p2ss0q-mlw &gt; Datastores

## Datastores

+ Create  Refresh  Unregister  Set as default datastore  Edit columns  Reset view

Showing 1-6 of 6 datastores

Name	☆	Type	Storage name
researcher_data		Azure Data Lake Storage Gen2	forskerplp2ss0qdls01
sp_data		Azure Data Lake Storage Gen2	forskerplp2ss0qdls01
workspaceworkingdirectory		Azure file share	forskerplp2ss0qstg
workspaceartifactstore		Azure Blob Storage	forskerplp2ss0qstg
workspacefilestore		Azure file share	forskerplp2ss0qstg
workspaceblobstore (Default)		Azure Blob Storage	forskerplp2ss0qstg



Region Hovedstaden

+ New

Home

Author

Notebooks

Automated ML

Designer

Assets

Data

Jobs

Components

Pipelines

Environments

Models

Endpoints

Manage

Compute

Datstores

Linked Services

Data Labeling

Region Hovedstaden &gt; forskerpl-p2ss0q-mlw

## forskerpl-p2ss0q-mlw



Create new ▾



## Notebooks

Code with Python SDK and run sample experiments.

Start now



## Automated ML

Automatically train and tune a model using a target metric.

Start now



## Designer

Drag-and-drop interface from prepping data to deploying models.

Start now

## Recent resources

Jobs

Compute

Models

Data

Display name



Experiment

Status

Logs

Submitted time

Submitte...

Job type



No jobs to display



## Notebooks

Files

Samples



Users

anders.christian.riis-jensen.01

fusion.yaml

globals.yaml

input\_tabular.yaml

output.yaml

test\_model.ipynb

test\_table.csv

SampleCode

data\_analysis.ipynb



Edit in VS Code (pr...

Compute instance:

No computes found

Your document is currently not connected to a compute. Switch to a running compute or create a new compute to run a cell.

Viewing

Last saved a few seconds ago

```
2 import azureml.core
3 from azureml.core.workspace import Workspace
4 from azureml.core import ScriptRunConfig, Environment, Experiment
5 from azureml.core.environment import CondaDependencies
6 from azureml.core import Workspace, Datastore, Dataset
7
8 # Initiate workspace
9 workspace = Workspace.from_config()
10
11 # Define datastore and load dataset
12 datastore_name = 'sp_data'
13 datastore = Datastore.get(workspace, datastore_name)
```

[1]



+ Code + Markdown



```
1 # Import libraries
2 import numpy as np
3 import pandas as pd
4 from datetime import datetime, date, timedelta
5 import matplotlib.pyplot as plt
6
7 pd.set_option('display.max_columns', None) # Allow to see all columns' content by scrolling
8 pd.set_option('display.max_rows', 40)
```

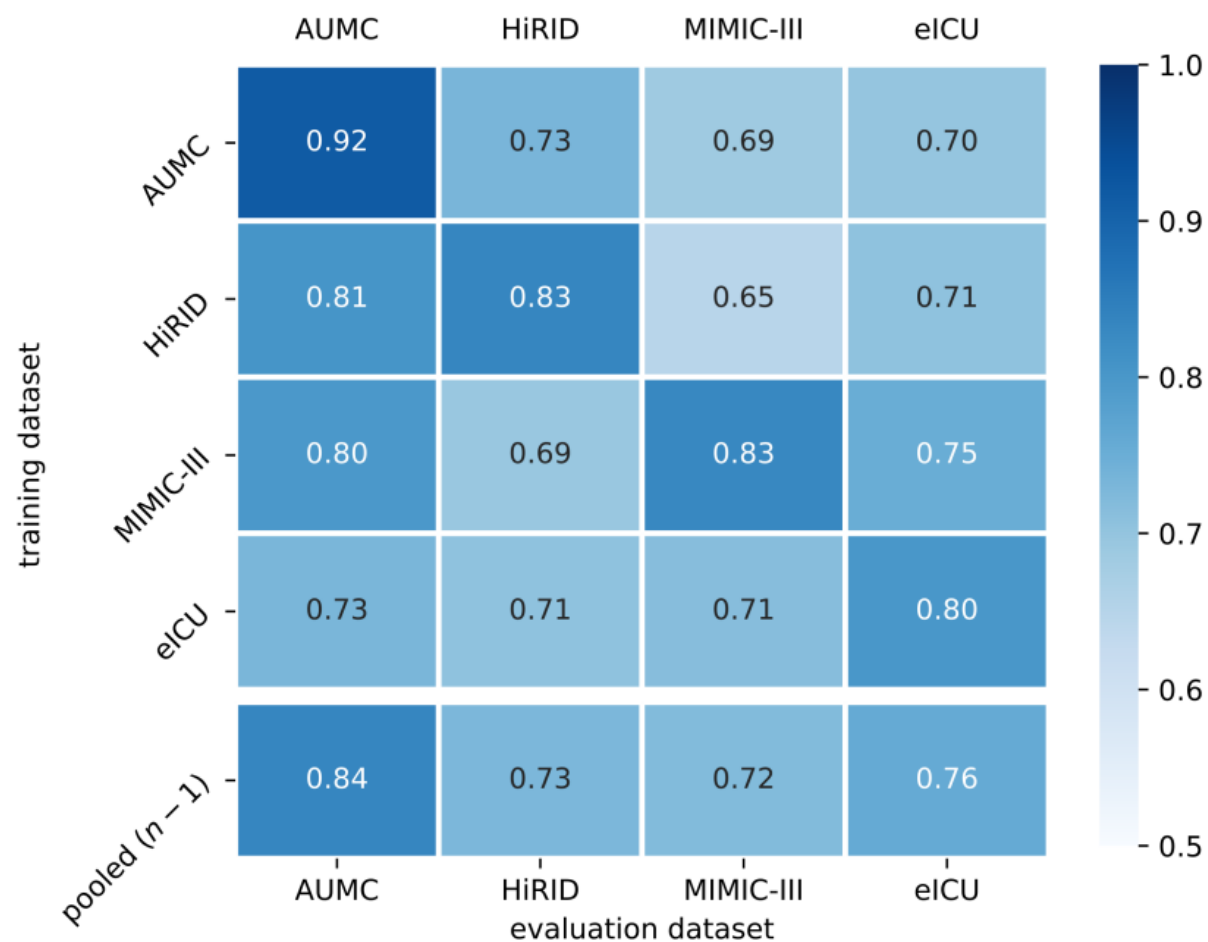
[2]



+ Code + Markdown

```
1 def col_count(df):
2     for col in df.columns:
3         print("-----", col, "-----")
4         print(df[col].value_counts())
```

## Performance af AI model på tværs af datasæt



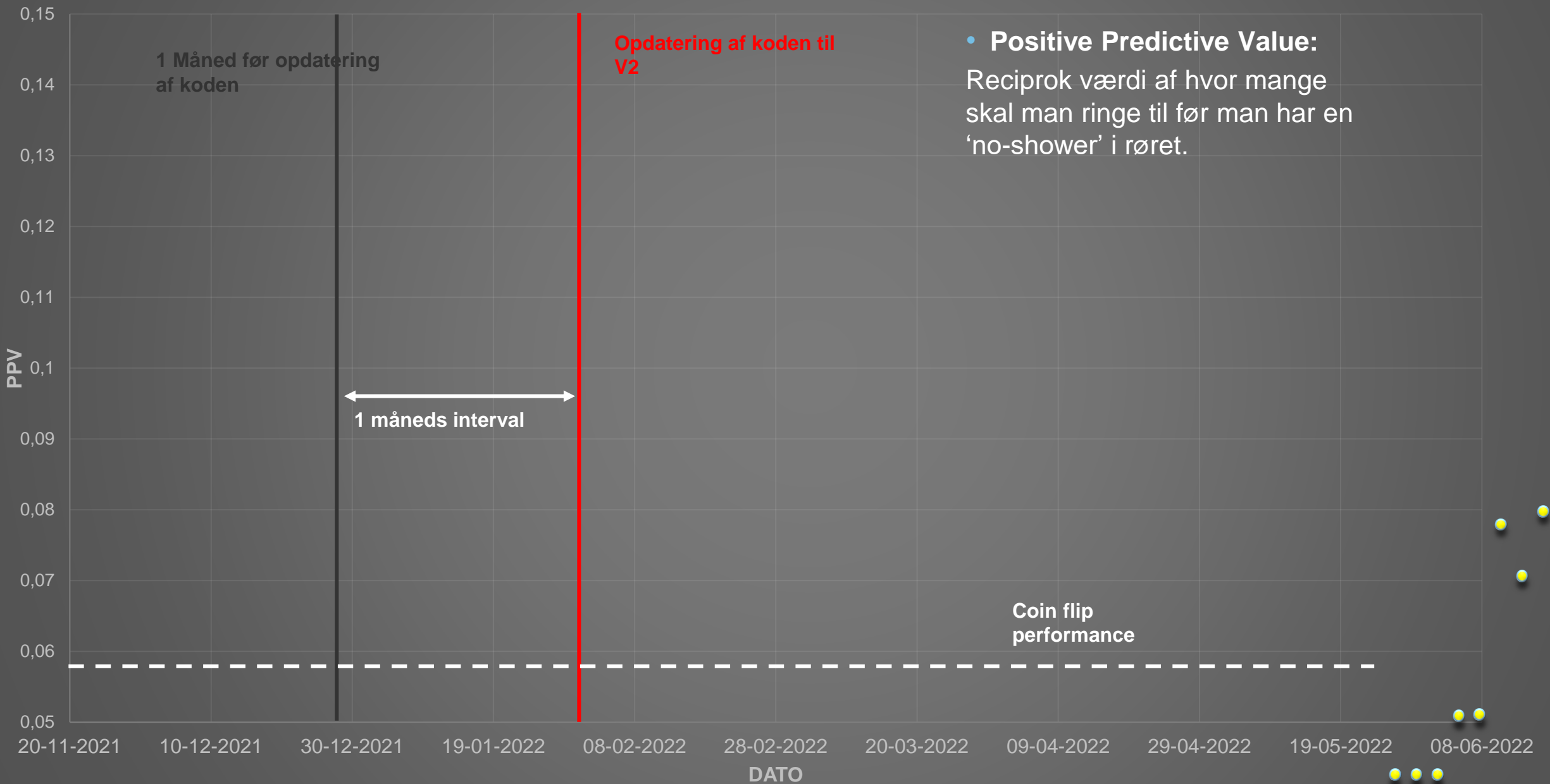
## AI modelmonitorering

- Definition af performance
- Definition af target og tidsskala kan gøre monitorering kompleks
- Ansvar for gen-træning



# Positive Predictive Value, BFH No-Show model

## 1 måneds sliding PPV beregning



## Tak for jeres opmærksomhed

- Spørgsmål?

Kontakt: [anders.christian.riis-jensen.01@regionh.dk](mailto:anders.christian.riis-jensen.01@regionh.dk)

