# Intro

* Dagens program / formål
* I holder selv pauser efter behov, vi holder opsamling kl. 11 (ca.)

# Opgave 1 Lidt om arrays, tensorer, slicing og sådan noget…

**Arbejd sammen i par (eller 3 personer, hvis det ikke går op…)**

Forklar begreberne

* + Skalar
  + Vektor
  + Matrice
  + NumPy array

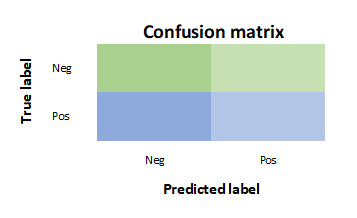
Løs følgende opgaver med papir og blyant (brug ***ikke*** computer!)

**Antag** at følgende tensor er defineret: numbers : [10, 45, 20, 67, 29, 12, 15, 99, 40, 23]

Løs følgende opgaver ved at finde et Python udtryk der giver korrekt svar (se løsning til 1. opg. for eksempel):

1. first\_number = numbers[0]
2. last\_number = ?
3. before\_last\_number = ?
4. seventh\_number = ?
5. first\_five\_numbers = ?
6. last\_three\_numbers = ?
7. third\_to\_sixth\_numbers = ?
8. reversed\_numbers = ?

# Opgave 2 Refleksionsspørgsmål

1. Hvad kan/skal man putte i en *pipeline* og hvad er formålet med en pipeline?
2. Hvilke metoder bliver kaldt på en pipelines komponenter under *training*?
3. Hvilke metoder bliver kaldt på en pipelines komponenter under *prediction*?
4. Hvad er fordelen ved både at have et *test* og *validation* dataset?
5. Beskriv *k-fold cross-validation* kort. Benytter den et validation dataset?
6. Hvad er *learning + validation curves* og hvad benyttes de til?
7. Hvad er *bias* og *variance* og hvilke problemer hører sammen med dem?
8. Hvad er *grid-search*?
9. Hvad er *performance evaluation metrics*?
10. Udfyld følgende *confusion matrix* med betegnelserne TP, FP, FN, TN (*true positive, false positive, false negative* og *true negative*) i følgende figur:
11. Hvis ovenstående tabel angår at afgøre retssager hvor man hellere vil lade en skyldig slippe fejlagtigt end at dømme en uskyldig, for hvilken metric vil du så optimere?  
    Hvis det derimod angår hvilke patienter du vil indkalde til nærmere undersøgelse og du hellere vil indkalde en rask end at misse en mulig syg patient, for hvilken metric vil du så optimere for?
12. Et billede, der indeholder tekst, himmel, orange, farverig

    Automatisk genereret beskrivelseBeregn ***ERR***, ***ACC***, ***FPR*,** ***TPR (REC)*, *PRE*** og ***F1*** for følgende confusion matrix:

# Opgave 3 Hands on pipelines og validering

Benyt dit dataset fra uge 7 (vedr. færgedata) og byg en pipeline som kører (!). Eksperimenter løs – fx med forskellige *transformers* og/eller *estimators*.

Benyt *cross validation* og udarbejd *training* samt *validation* kurver.

Prøv med et nyt dataset fra *Kaggle*.

# Opgave 4 … lidt mere om tensorer

**Antag** nu at følgende NumPy array er defineret: numbers = np.array ([

[0, 1, 2],

[10, 11, 12],

[20, 21, 22],

[30, 31, 32],

[40, 41, 42] ])

* # Find count of rows and colums:  
   rows\_count, columns\_count = ?, ?
* # Get the first element of each row and save it into array with shape (5,):  
   first\_elems\_A = ?
* # Get the first element of each row and save it as 2-dimensional array (with shape (5, 1)):  
   first\_elems\_B = ?
* # Get the second and third elements from the fourth and fifth rows:  
   some\_elems = ?
* # Divide each number by 2:  
   numbers\_divided\_by\_two = ?
* # Add 5 to each number:  
   numbers\_plus\_five = ?
* # Add this array: [[5, 6, 7], [15, 16, 17], [25, 26, 27], [35, 36, 37], [45, 46, 47]] to the "numbers" array:  
   new\_numbers = ?

# Opsamling / Outro (5 minutter)

* Var der noget vi ikke nåede eller ikke fik behandlet tilstrækkeligt?
* Var det svært / nemt?
* Lærte du tilstrækkeligt i dag?   
  Hvis nej: Hvad var årsagen?
* Andet og næste gang…