

# Startvergadering ML sessie 6

Team  $\exists$ uler

*Daan, Marie, Zeineb, Florian, Vincent, Jasper, Lasha, Younes*

Vrijdag 17 november 2023

## Stand van zaken

Afgelopen week werden alle simulaties en procedures van vorige sessie afgewerkt. In figuur 1 is de status van het tabblad *Implementatie* te zien vóór de aanvang van deze 6e sessie.

Bovendien programmeerde Vincent al een aantal zaken om voor de bijkomende toepassing *Support Vector Machines*. Hij maakte eerst een SVM-model op een trainingsdataset met 30 onafhankelijke variabelen of x-waarden. Dit model kon zeer accuraat voorspellen of een tumor in de borst van een persoon al dan niet kwaadaardig was bij een validatie- en testdataset. Het probleem met het gebruiken van zo veel inputs, is echter dat we dit niet visueel kunnen voorstellen aangezien we in heel hoge dimensies werken. Daarom werd er ook een bestand aangemaakt waarin er slechts 2 inputs worden bekeken, zodat we de data mooi kunnen plotten. Dit alles kan teruggevonden worden in de subdirectory *tumorclassificatie* in de map *Implementatie*.

## Planning voor deze sessie

### Bespreken van de simulaties en procedures

Deze sessie zullen we de simulaties en procedures van vorige sessie eerst en vooral met elkaar bespreken, zodat iedereen zeker op de hoogte is van wat er precies allemaal in de Python-files en Jupyter-notebooks staat. Het is belangrijk dat iedereen de geschreven code snapt. Op die manier kunnen we allemaal samen een coherent eindverslag maken en heeft iedereen genoeg kennis om tijdens de postersessie een woordje uitleg te kunnen geven.

### Uitwerken van de extra toepassing

Na het bespreken van de simulaties uit vorige sessie, zullen we ook eerst de toepassing rond de classificatie tumoren bespreken. Daarna zullen we voor deze toepassing simulaties uitvoeren, waarbij we telkens maar zullen kijken naar 2 inputs, aangezien we anders de data niet grafisch kunnen voorstellen. De zaken die in orde moeten worden gebracht, staan reeds in de KanBan en zijn te zien in figuur 2. Ieder teamlid kiest dus een taak uit deze lijst.

Te doen (type+naam)	Bezig	Afgewerkt	Bestand	Finaal	Naam 1
Procedure xsample(n,a,b)		Vincent	procedures.py	OK	
Voorbeelden xsample(n,a,b)		Vincent	voorbeelden.ipynb	OK	
Procedure ysample(x,f,e)		Vincent	procedures.py	OK	
Voorbeelden ysample(x,f,e)		Vincent	voorbeelden.ipynb	OK	
Procedure ols(y,x)		Vincent	procedures.py	OK	
Voorbeelden ols(y,x)		Vincent	voorbeelden.ipynb	OK	
Procedure knn(x0,y,x,k)		Vincent	procedures.py	OK	
Voorbeelden knn(x0,y,x,k)		Vincent	voorbeelden.ipynb	OK	
Procedure mknn(x0,y,x,k)		Vincent	procedures.py	OK	
Voorbeelden mknn(x0,y,x,k)		Vincent	voorbeelden.ipynb	OK	
Simulatie 1		Vincent	simulaties.ipynb, illustraties/simulatie1.pdf	OK	
Simulatie 2		Vincent	simulaties.ipynb, illustraties/simulatie2.pdf	OK	Daan niet afgewerkt
Simulatie 3		Vincent	simulaties.ipynb, illustraties/simulatie3.pdf	OK	Marie niet afgewerkt
Procedure mls(y,X)		Vincent	procedures.py	OK	
Voorbeelden mls(y,X)		Vincent	voorbeelden.ipynb	OK	
Simulatie 4		Vincent	simulaties.ipynb, illustraties/simulatie4.pdf	OK	
simulatie 5		Vincent	simulaties.ipynb, illustraties/simulatie5.pdf	OK	Florian niet helemaal juist
Simulatie 6		Vincent	simulaties.ipynb, illustraties/simulatie6.pdf	OK	Lasha niet afgewerkt
Simulatie 7		Vincent	simulaties.ipynb, illustraties/simulatie7.pdf	OK	
Simulatie 8		Vincent	simulaties.ipynb, illustraties/simulatie8.pdf	OK	Zeineb niet afgewerkt
Procedure berekenKOpt		Vincent	procedures.py	OK	
Voorbeelden berekenKOpt		Vincent	voorbeelden.ipynb	OK	
Procedure berekenPOpt		Vincent	procedures.py	OK	
Voorbeelden berekenPOpt		Vincent	voorbeelden.ipynb	OK	
Simulatie 9		Vincent	simulaties.ipynb, illustraties/simulatie9.pdf, il	nakijken	

Figuur 1: De status van het tabblad *Implementatie* in de KanBan vóór de aanvang van sessie 6.

Te doen (type+naam)
Simulatie 1: vergelijking tussen C-waarden bij SVM met 2 inputs
Illustratie: plot de grafieken voor een 3 verschillende waarden van C naast elkaar met subplot
Illustratie: plot de accuraatheid ifv C
Simulatie 2: vergelijking tussen kernels bij SVM met 2 inputs
Illustratie: plot de grafieken voor de 4 verschillende kernels naast elkaar met subplot
Illustratie: plot de accuraatheid ifv de kernel
Simulatie 3: vergelijking tussen gamma-waarden bij SVM met 2 inputs
Illustratie: plot de grafieken voor 3 verschillende waarden van gamma naast elkaar met subplot
Illustratie: plot de accuraatheid ifv gamma
Simulatie 4: bekijk hoe het SVM model beïnvloed wordt door het kiezen van twee andere onafhankelijke variabelen (x-waarden)
Procedure voor het berekenen van de optimale metaparameters (3) gegeven een trainingset en een validatieset

Figuur 2: De planning voor sessie 6 in het tabblad *Toepassing*.