Maskinlæring – forløb (del 2)

Gruppenavn:

**Datarepræsentationer**

Gå til [kortlink.dk/2qscq](http://kortlink.dk/2qscq) og eksperimenter med det forskellige datarepræsentationer.

Lav en model med tre klasser: Stille og 2 I selv vælger. Lav 5-10 eksempler for hver klasse. Forsøg at forbedre modellen og udfyld skemaet nedenfor:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Klasser |  (Fx cirkel) |   |   |  |
| Antal eksempler  |   |   |   |  |
| Vælg to datarepræsentationer.Skriv hvilke i valgte og hvorfor. |  |
| Hvad gjorde i for at forbedre modellen?  |   |

Nu skal i prøve at tegne datapunkter ind i et koordinatsystem i hånden.

Vælg en akse (x, y eller Z) og kig på de to datarepræsentationer i valgte ovenfor. Tegn et koordinatsystem og start med at putte datapunkterne for stille og en af det andre klasser ind.

Kig på punkterne. Kan i lave (tegne) en model, der kan klassificere mellem de to klasser, når der kommer nye datapunkter ind?

Tegn punkterne ind for den tredje klasse. Kan i lave (tegne) en model, der kan klassificere mellem alle tre klasse?

Den model I har tegnet arbejder i 2 dimensioner. Hvad ville I gøre, hvis I skulle arbejde i tre dimensioner?

Hvor mange dimensioner bruger modellen i værktøjet, når alle datarepræsentationer er slået til? Prøv at kigge på visualiseringerne af datarepræsentationerne i værktøjet, hvis I er i tvivl.

**K-nærmeste naboer**

Gå til <https://v3-2.ml-machine.r59q.com/> og eksperimenter med den nye KNN model. Visualiseringen er ikke perfekt, men undersøg modellen.

Lav igen en model med tre klasser: Stille og to i selv vælger. Lav 8-10 eksempler for hver klasse. Vælg to data repræsentationer.

Undersøg modellen. Hvordan tror i den virker?

I kan ændre k-værdien, prøv at lege med det og besvar:

* Hvad sker der, når k er sat til 1?
* Hvad sker der, når k er sat til 8?

Hvad sker der, når i vælger flere datarepræsentationer (og visualiseringen forsvinder)?

Hvilken type matematik tror I bliver brugt i en k-nærmeste naboer model? Hvorfor?

Hvordan ville I beregne afstanden mellem to punkter?

**Opsamling**

*Tag gerne udgangspunkt i jeres indledende refleksioner, og overvej om I ville omformulere jeres forklaringer fra del 1.*

Forklar med egne ord hvad maskinlæring og kunstig intelligens er og kan?

 Tænk på den artikel, som I læste. Hvordan vil I lave en model for predictive policing?

* Hvilken data vil I indsamle? Hvorhenne
* Hvilken datarepræsentationer vil I bruge?
* Hvad vil i være særligt opmærksomme på kunne gå galt?
* Hvad vil i gøre for at undgå, at det går galt?