Machine Learning / Deep Learning Forløb

Hvis en af de følgende links ikke længere er gyldigt venligst sig det til læreren.

# Afsnit 1: Machine Learning

## ****Kom i gang****

**Se denne video for at få et overblik over de mange forskellige termer og underdiscipliner inden for AI. Denne video kan være lidt overvældende, hvis du næsten ikke ved noget om maskinlæring/AI. Det er bedst at se videoen igen, når du har arbejdet dig igennem de følgende tutorials. Så vil du være i stand til at kategorisere mange af de termer, der nævnes i videoen. Det vil give dig en bedre forståelse af, hvilke områder af maskinlæring/AI disse tutorials dækker, og hvad der ellers er at opdage.**

<https://www.youtube.com/watch?v=Fa_V9fP2tpU>

**Gennemfør disse tutorials:**

1. <https://www.kaggle.com/learn/pandas>
2. <https://www.kaggle.com/learn/data-cleaning>
3. <https://www.kaggle.com/learn/data-visualization>
4. <https://www.kaggle.com/learn/intro-to-machine-learning>
5. <https://www.kaggle.com/learn/intermediate-machine-learning>
6. <https://www.kaggle.com/learn/feature-engineering>

## Alternativ tilgang til at komme i gang med maskinlæring

Opret en konto på kaggle.com

Deltag i denne konkurrence: <https://www.kaggle.com/competitions/titanic>

Følg instruktionerne på denne side.

Få et overblik over strukturen og indholdet af Titanic-træningsdataene. Du kan f.eks. bruge bibliotekerne matplotlib og måske også seaborn. Det kan f.eks. være interessant at se på, hvilke af attributterne der er stærkt relaterede. (Nøgleord: korrelationer).

For at læse og manipulere data skal du bruge pandas-biblioteket og sandsynligvis også numpy.

Væn dig til arbejdsmiljøet i Kaggle ved at oprette en første notesbog, som du bruger til at indsende et første svar til konkurrencen. Det gør ikke noget, hvis du kun får en score på 0 % i første omgang.

Find algoritmer til at forbedre din forudsigelsesmodel baseret på træningsdata.

Opnå en score på over 77% i konkurrencen! (Alle deltagere med en score på 100 % er kedelige snydere).

Der er gode vejledninger til matplotlib, seaborn, pandas og numpy på <https://www.kaggle.com/learn>. Der finder du også instruktioner om, hvordan du træner modeller med træningsdata.

Du kan selvfølgelig også bruge den officielle dokumentation for disse biblioteker eller Google eller chatbots. <https://www.perplexity.ai/> er en god chatbot, fordi den giver kilder.

## ****Projektbeskrivelse****

**Deltag i denne konkurrence:**

<https://www.kaggle.com/c/house-prices-advanced-regression-techniques>

Konkurrencen ligner meget det, du har arbejdet med i ovennævnte vejledninger. Prøv at klatre op i ranglisten så vidt du kan.

Derefter vælg en anden kaggle-konkurrence i samråd med din lærer.