Projekt: Predicting Optimal Fertilizer

### Projektmål

Lær at udvikle en ”predictive model” baseret på en datatabel.

Deltag i konkurrencen [Predicting Optimal Fertilizers](https://www.kaggle.com/competitions/playground-series-s5e6/) på kaggle.com til dette formål.

Som altid gælder ved hvert trin i projektet: Tal med din lærer, hvis du sidder fast eller er usikker på, hvad du skal gøre som det næste.

### 1. Læs konkurrencebeskrivelsen

Få først et overblik. Du behøver ikke læse eller forstå alt lige nu.

### 2. Forstå datastrukturen

Forstå betydningen af alle kolonner i datatabellen. Gå til siden med den [oprindelige tabel](https://www.kaggle.com/datasets/irakozekelly/fertilizer-prediction) for at få hjælp.

Konkurrencen bruger en tabel med kunstigt ændrede data, men med samme struktur. Tabellen er ikke særligt godt dokumenteret på kaggle.com. Find andre informationskilder eller lav dine egne antagelser. Overvej hvor vigtigt det er at forstå betydningen af de enkelte kolonner korrekt. Tal med din lærer om det!

### 3. Forstå konkurrencevurderingsfunktionen

Vurderingsfunktionen "Mean Average Precision @ 3" forklares dårligt på kaggle.com. Find en bedre forklaring. Hvor mange point får du pr. række, hvis du nævner den rigtige gødning på første, anden eller tredje plads eller slet ikke? Tal med din lærer om det!

### 4. Forstå det format, du skal aflevere din løsning i.

Læs kravene til formatet for filen sample\_submission.csv.

### 5. Lav en første test

Skriv kode, der indlæser test- og træningsdata og genererer sample\_submission.csv. Først kan du bare vælge 3 tilfældige gødningstyper, som du skriver i hver række i din løsning. Indsend løsningen. Så længe du får en score over 0, var testen vellykket.

### 6. Tilføj din første model

Træn nu en model med træningsdataene, og anvend modellen på testdataene for at generere din løsning. Indsend løsningen. Du kan evt. genbruge noget af din kode fra Titanic-opgaven.

### 7. Optimer din forudsigelse

Eksperimenter med parametrene i din model. Prøv forskellige modeller. Kombinér flere modeller.

Meget vigtigt: **Feature Enhancement**. Forvandl dataene på en meningsfuld måde, eller skab nye datakolonner ud fra de eksisterende data. Fjern eventuelt nogle datakolonner.

Inkluder den oprindelige tabel.

### 8. Lad dig inspirere

Se på [andres kommenterede løsninger](https://www.kaggle.com/competitions/playground-series-s5e6/discussion?sort=hotness). Eller tag et kig på den [bedste løsning](https://www.kaggle.com/competitions/playground-series-s5e6/writeups/chris-deotte-1st-place-fast-gpu-experimentation-wi). Eller kig på konkrete [kodeeksempler](https://www.kaggle.com/competitions/playground-series-s5e6/code?competitionId=91717&sortBy=voteCount&excludeNonAccessedDatasources=true) Brug kun kode fra andre, hvis du forstår, hvad den gør! Det er meget muligt, at du skal undersøge tingene grundigt og lære nye koncepter, før du kan bruge andres løsninger.

Søg efter optimeringer, indtil du ikke længere kan forbedre din score nævneværdigt.

### 9. Skriv din løsning ned

Post en notebook af din løsning under afsnittet "Discussion" i konkurrencen. Gør en stor indsats for at forklare din kode, dit feature enhancement og din tilgang. Så kan et link til denne notebook være en værdifuld berigelse til dine fremtidige jobansøgninger.