

Aktivitetskort: APHACA-maskinen

Introduktion: Hvordan genererer computere tekst?

Du skal nu undersøge, hvordan en computer kan generere ny APHACA lyrik, v.h.a. et datasæt, og hvordan sprogmodeller balancerer kreativitet og troværdighed.

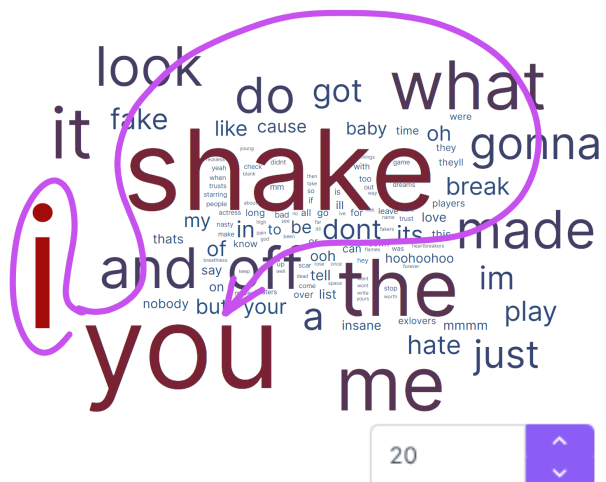
- Start her: [Hvordan kan en computer forstå og generere tekst?](#)
- Vælg datasættet “APHACA” og START AKTIVITETEN.

1 Ordsky

En af de enkleste måder at generere tekst på, er at tage det mest almindelige første ord i hele datasættet, derefter det mest almindelige andet ord, derefter det tredje, og så videre. Fx “i shake ...” hos Taylor Swift

Rens din tekst: Skru op for, hvor mange gange et ord skal forekomme, for at det er med i din ordsky.

Undersøg: Kan du lave en APHACA-sætning ved at forbinde det største, det næststørste og det tredjestørste ord fra din ordsky? (Skriv din sætning i boksen)



Det mest almindelige ord	Det næstmest almindelige	Det tredje mest almindelige

Rens videre: Aktiver STRAMNING af danske ord og filtrer danske STOPORD:

Undersøg: Kan I lave en sætning ved at forbinde det største, det næststørste, det tredjestørste og det fjerdestørste ord? Hvis ja, så skriv sætningen i boksen.

Det mest almindelige ord	Det næstmest almindelige ord	Det tredje mest almindelige ord	Det fjerde mest almindelige ord

Positions-generering minder om det, du lige har forsøgt, med de "største" ord i ordskyen. Nu skal du prøve det i maskinrummet.

2

Positions-tekst-generering

Positions-generering er den enkleste måde at begynde at generere tekst på.

Grådig ▾

Generer

Hvordan?

- **Klik generer** og skab en nye sætninger
- **Skift temperatur** mellem *Grådig*, *Vægtet tilfældigt* og *Tilfældig*.
- **Undersøg** hvordan temperaturen påvirker modellens kreativitet og troværdighed.
- **Klik på "hvordan"** og få vist de ord, der også kunne være på position 1, 2, 3
 - ☐ Hvilken temperatur er bedst til at lave nye sætninger (Grådig, Vægtet tilfældigt, Tilfældig)?
 - ☐ Hvilken temperatur er bedst til at lave meningsfulde sætninger med (Grådig, Vægtet tilfældigt, Tilfældig)?
 - ☐ Er der en indstilling, som giver nye og troværdige sætninger (Grådig, Vægtet tilfældigt, Tilfældig)?
 - ☐ Hvad er den længste række af ord, som kan få til at give mening?

3

N-grams: Tager du hensyn til konteksten?

Positions-genereringen fungerer godt i starten, men længere inde i sætningen, bliver sandsynlighederne og positionerne mere rodede. Derfor kan det give mening at se på, hvilke ord der optræder sammen, i stedet for hvornår de optræder. N-grammer i en ordsky er en måde at gøre dette på. Skyen viser grupper af ord, der optræder i teksten. Hvordan ser N-grammerne ud i dit datasæt og kan de kobles sammen til en sætning?

- **Aktivitet:** Lav en ordsky med N-grammer?

brug for
du er jeg har

4

N-grams tekst-generering

Kan du skrive 10 sætninger som både er nye og troværdige APHACA-linjer?

- Generer sætninger ved at justere på n-gram og sandsynlighedsmodellen
- Skriv de nye sætninger ned og brug HVORDAN-knappen til at undersøge sandsynligheden for andre ord. Noter et eller to af disse alternative ord. Ville de gøre sætningen bedre, mere nyskabende eller mere APHACA-roværdig?
- Skriv også interessante fejl eller halv-gode sætninger ned.

2-gram:

Nedenfor bruger vi N-gram

Mere sands

Mindre sandsynligt

magt

Position: 4

i min ⇒ pude

mig i min pude [SLUT]

Himmel

✓ Dette er en ny sætning, der ikke findes i datasættet

Hvordan?

mig	i	min	pude	[SLUT]		
			himmel/magt			

[illegible]