

Danish Foundation Models: Sprogmodeller til dansk tale og tekst

Store sprogmodeller, som bl.a. ligger til grund for ChatGPT, har transformeret adskillige forskningsområder. Dog risikerer mindre sprog som Dansk at sakke bagud på grund af høje træningsomkostninger og små incitamenter for store virksomheder til at træne disse modeller. For at bekæmpe dette søger the Danish Foundation Models projektet at levere og validere åbne og høj kvalitets sprogmodeller for det danske sprog. Dette opnås gennem bredt samarbejde med offentlige og private institutioner, for at sikre høj datakvalitet og anvendelighed af de trænede modeller. Vi præsenterer projektets motivation, den aktuelle status og fremtidige perspektiver.