

ONDERZOEKSVOORSTEL

Een applicatie die dagelijks een AI-gegenereerde kunstwerk deelt op basis van een automatische analyse van nieuwswebsites of socialmediaplatformen: Toegepast onderzoek.

Bachelorproef, 2022-2023

Dario Bronders

E-mail: dario.bronders@student.hogent.be

Co-promotor: M. De Buck (We-Are, manu@we-are.be)

Samenvatting

In dit toegepast onderzoek wordt er in de eerste fase onderzocht op welke manier we nieuwswebsites of social media platformen kunnen scrapen aan de hand van een python library BeautifulSoup. Deze zal dan ook worden geïmplementeerd. In de tweede fase gaan we onderzoeken hoe we deze data kunnen rangschikken en filteren, op deze manier achterhalen we de kernzaak van de dag. Eenmaal we dit achterhaald hebben, kunnen we dit vervolgens gebruiken om een kunstwerk te genereren op basis van de tekstuele input. Hiervoor zullen we deep learning models op toepassen die in 2022 publiek beschikbaar zijn gesteld.

Keuzerichting: Mobile & Enterprise development

Sleutelwoorden: Data scrapen en analyseren, Stable Diffusion en Dall-E, AI-gegenereerd kunstwerk

Inhoudsopgave

1	Introductie	1
2	Literatuurstudie	1
3	Methodologie	1
4	Verwacht resultaten	2
	Referenties	2

1. Introductie

De meestgekende nieuwsbronnen proberen al jaren objectief en feitelijk te blijven om informatie vanop eenzelfde standpunt en met een gelijkaardig boodschap over te brengen.

Binnen mijn toegepast onderzoek zal ik een toepassing maken die een kunstwerk zal genereren met behulp van één of meerdere deep learning modellen. Dit zou er voor zorgen dat het dagelijkse hoogtepunt geabstraheerd kan worden tot een unieke AI-gegenereerde kunstwerk.

2. Literatuurstudie

(Salem & Mazzara, 2020) (Lauri, 2019)

3. Methodologie

Inleiding

Het toegepast onderzoek begint 2 maart 2023 en zal beëindigd worden voor 28 mei 2023.

Fase 1: Realiseren van een scraper

Om de data te bekomen van de verschillende

soorten websites of social-media platformen zal er een web scraper worden gemaakt. Deze scraper zal ontwikkeld worden in python met behulp van een externe library *BeautifulSoup*.

Doordat de presentatie van de verschillende artikelen kunnen verschillen in taal en structuur, zal de scraper een algoritme implementeren die het mogelijk maakt om op een uniforme manier verschillende websites te scrapen.

Fase 2: Data verwerken en analyseren

Tijdens de tweede fase zullen we onderzoeken op welke manier we de bekomen data uit voorgaande fase kunnen analyseren en sorteren.

Het zal belangrijk zijn om rekening te houden met de volgende vragen:

- Wat zijn de te extraheren kernzaken?
- Wat is het sentiment van de dag?
- Welke topic komt het vaakst voor?
- Op basis van welke gegevens kunnen we de artikels sorteren?

Nadat er een gepaste methode wordt gevonden om dit te realiseren, zal deze ook geïmplementeerd worden. Op deze manier kunnen we steeds het belangrijkste artikel van de dag eruit halen.

Fase 3: Kunstwerk genereren

Nu dat we weten uit de vorige fase wat het hoogtepunt van de dag was. Kunnen we hierop een

kunstwerk laten genereren.

Hiervoor zal er gebruik gemaakt worden van een of meerdere deep learning modellen Dall-E en/of Stable Diffusion die de kerntekst van een artikel zal omvormen tot een foto.

Één of beide technologieën zullen gebruikt worden.

4. Verwacht resultaten

Een applicatie ontwerpen die dagelijks een kunstwerk kan genereren op basis van het hoogtepunt van de dag.

Op 27 oktober, toen Marokko won van België tijdens de WK kon het hoogtepunt in België *'Riots in Brussels after soccer game, painting'* geweest zijn. Hieronder vindt u enkele voorbeelden die gegenereerd zijn met behulp van Dall-E op basis van deze tekstininput.



Referenties

- Lauri, K. (2019). *A web scraping system for extracting news articles*. (masterscriptie). Aalto University.
- Salem, H., & Mazzara, M. (2020). Pattern Matching-based scraping of news websites. *Journal of Physics: Conference Series*, 1694, 6.